

[Save Link](#) Saves this search as a Durable Link under "Results-Marked List"

At least 171 articles matched your search.

- ☐ 161. [Archiving the Internet / Brewster Kahle makes digital snapshots of Web; \[FINAL Edition\]; Carolyn Said, Chronicle Technology Editor, San Francisco Chronicle, San Francisco, Calif.; May 7, 1998; pg. B.3](#)
- ☐ 162. [Corbis Acquires Westlight; Corbis Expands its Product Spectrum with the Purchase of Premier Commercial Stock Photography Agency; Business Editors; Business Wire, New York; May 5, 1998; pg. 1](#)
- ☐ 163. [SIEMENS NIXDORF: Siemens enables TV and Internet on the PC at no extra cost, thanks to ZDF InterCast; M2 Presswire, Coventry; May 5, 1998; pg. 1](#)
- ☐ 164. [Rearden Technology Ships SiteCam 2.0, Now with Dial-up Support; Full-Featured WebCam Software Now with Advanced Anti-Hijacking and Support for Multiple Video Digitizers; Business Editors; Business Wire, New York; May 4, 1998; pg. 1](#)
- ☐ 165. [GERMANY: COMPUTER SOFTWARE MARKET \(2\); Industry Sector Analysis, New York; May 2, 1998; pg. 1](#)
- ☐ 166. [Server solutions on NT or Unix; Anonymous; American Printer, Chicago; May 1998; Vol. 221, Iss. 2; pg. 84, 1 pgs](#)
- ☐ 167. [Online Journals: Moving in the Right Direction; Marc S Rovner, Chest, Chicago; May 1998; Vol. 113, Iss. 5; pg. 1423, 1 pgs](#)
- ☐ 168. [This Is BQ's Best of All Possible Worlds; Barbara Quint; Information Today, Medford; May 1998; Vol. 15, Iss. 5; pg. 8, 3 pgs](#)
- ☐ 169. [S.H.O.W.M.E.: Spear-heading online work in mathematics education; Gene Abrams; T.H.E. Journal, Tustin; May 1998; Vol. 25, Iss. 10; pg. 53, 3 pgs](#)
- ☐ 170. [Attachmate releases new DataBridge NT Clients V2.5; Anonymous; Unisys World, Austin; May 1998; Vol. 19, Iss. 5; pg. 17, 1 pgs](#)

[First](#) [Previous](#) [Next](#) [Last](#)

151-160 171-180

Refine your search. Enter a word, words or specific phrase.

Date range:

Publication type:

Search in:

- ☒ Show results with full text availability only
- ☐ Show articles from peer reviewed publications only
- ☐ Show total number of articles

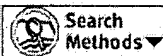
ProQuest®

[Return to NPL Web Page](#)[Text Version](#)

English

[?Help](#)

Collections



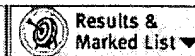
Search Methods



Topic Finder



Browse Lists



Results &amp; Marked List



Search Guide

Searching collections: All Collections

Search Results

[Save Link](#) Saves this search as a Durable Link under "Results-Marked List"**At least 11 articles matched your search.**

- ☐ 1. [Kopin and Sound Vision Announce Partnership to Provide Complete Digital Camera Reference Designs; Designs to Provide Total OEM Solution; \*Business/Technology Editors; Business Wire\*, New York; Dec 15, 1998; pg. 1](#)
- ☐ 2. [GOLDSBORO IS ONE BIG LABORATORY TROUBLED SCHOOL IS REBORNAS SHOWPIECE HIGH-TECH MAGNET; \[SEMINOLE Edition\]; \*Leslie Postal of The Sentinel Staff, Orlando Sentinel\*, Orlando, Fla.; Dec 13, 1998; pg. K.1](#)
- ☐ 3. [NurLogic Joins with IBM to Offer 0.25-Micron and 0.18-Micron Standard Cell Libraries; \*PR Newswire\*, New York; Dec 9, 1998; pg. 1](#)
- ☐ 4. [ROYAL PROJECTS: Digital library to be first for Thailand; \*The Bangkok Post\*, Bangkok; Dec 9, 1998; pg. 1](#)
- ☐ 5. [Sound Vision and IBM To Cooperate on New Digital Imaging Technology Companies To Focus on New Imaging Solutions; \*Business/Technology Editors; Business Wire\*, New York; Dec 7, 1998; pg. 1](#)
- ☐ 6. [Everything you wanted to know about digital imaging but were afraid to ask; \*Roy Tennant, Library Journal\*, New York; Dec 1998; Vol. 123, Iss. 20; pg. 33, 2 pgs](#)
- ☐ 7. [Information imagineering: Meeting at the interface; \*Louise Cartee; Library Software Review\*, Westport; Dec 1998; Vol. 17, Iss. 4; pg. 294, 1 pgs](#)
- ☐ 8. [The management of digital data: A metadata approach; \*Alison Chilvers; The Electronic Library\*, Oxford; Dec 1998; Vol. 16, Iss. 6; pg. 365, 8 pgs](#)
- ☐ 9. [Uniting islands of data; \*Herman Mehling; Computer Reseller News\*, Jericho; Nov 30, 1998, Iss. 819; pg. 95, 2 pgs](#)
- ☐ 10. [Netwatch; \*KARLIN LILLINGTON; The Guardian\*, Manchester \(UK\); Nov 26, 1998; pg. 012](#)

[Next](#)

11-20

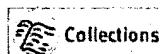
Refine your search. Enter a word, words or specific phrase.

[Search](#)Date range: Publication type: Search in: 

- ☒ Show results with full text availability only
- ☐ Show articles from peer reviewed publications only
- ☐ Show total number of articles

[Return to NPL Web Page](#)[Text Version](#)

English

[?Help](#)

Collections



Search Methods



Topic Finder



Browse Lists



Results &amp; Marked List



Search Guide

Searching collections: All Collections

Search Results

[Save Link](#) Saves this search as a Durable Link under "Results-Marked List"**At least 41 articles matched your search.**

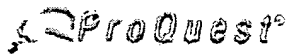
- ☐ 31. [6 firms vie to create 'wired city'; \[5XS Edition\]; SUSAN DODGE; Chicago Sun - Times](#), Chicago, Ill.; Sep 24, 1998; pg. 14
- ☐ 32. [Assembling the world's biggest library on your desktop; Joseph Alper, Science](#), Washington; Sep 18, 1998; Vol. 281, Iss. 5384; pg. 1784, 3 pgs
- ☐ 33. [Taming MEDLINE with concept spaces; Joseph Alper, Science](#), Washington; Sep 18, 1998; Vol. 281, Iss. 5384; pg. 1785, 1 pgs
- ☐ 34. [Long-distance learning; John Sharp; Fiscal Notes](#), Austin; Sep 1998; pg. 1, 4 pgs
- ☐ 35. [UMI's Digital Vault Initiative project; Peter Jacso; Information Today](#), Medford; Sep 1998; Vol. 15, Iss. 8; pg. 15, 2 pgs
- ☐ 36. [Building the digital library; Elana Varon; Federal Computer Week](#), Falls Church; Aug 24, 1998; Vol. 12, Iss. 29; pg. 1, 3 pgs
- ☐ 37. [Institute Brings Riches of Library of Congress to Hewlett-Annenberg Teachers and Librarians; Business Editors; Business Wire](#), New York; Aug 17, 1998; pg. 1
- ☐ 38. [Extensis Corporation and Digital Media, Inc. Announce Cooperative Development Of New Training Titles; PR Newswire](#), New York; Aug 5, 1998; pg. 1
- ☐ 39. [JUST FOR KIDS; WWW.4kids.org; \[FINAL Edition\]; The Sun](#), Baltimore, Md.; Aug 3, 1998; pg. 8.E
- ☐ 40. [Manage your metadata; Robert J Boeri; E Media Professional](#), Wilton; Aug 1998; Vol. 11, Iss. 8; pg. 42, 1 pgs

◀◀ First ◀ Previous Next ▶▶  
21-30 41-50

Refine your search. Enter a word, words or specific phrase.

 Date range: Publication type: Search in: 

- ☒ Show results with full text availability only
- ☐ Show articles from peer reviewed publications only
- ☐ Show total number of articles

[Return to NPL Web Page](#)[Text Version](#)

English

[?Help](#)

Collections

Search  
MethodsTopic  
FinderBrowse  
ListsResults &  
Marked ListSearch  
Guide

Searching collections: All Collections

Article Display

[Email Article](#)

Article 40 of 42

[Publisher Info.](#)[Print Article](#)☐ Mark article

Article format: Full Text

[Save Link](#)

Saves this document as a Durable Link under "Results-Marked List"

## Manage your metadata

*E Media Professional*; Wilton; Aug 1998; [Robert J Boeri](#); [Martin Hensel](#);**Sub Title:** The Magazine for Electronic Media Producers & Users**Volume:** 11**Issue:** 8**Start Page:** 42**ISSN:** 1090946X**Subject Terms:** [Information management](#)[Protocol](#)[Searches](#)[World Wide Web](#)[Internet](#)[Information management](#)[Interoperability](#)**Classification Codes:** 5240: *Software & systems*9000: *Short article***Geographic Names:** US**Abstract:**

ICE, or Information Content and Exchange, is a protocol aiming to develop a consistent vocabulary for describing and managing the exchange of content and electronic assets between Web businesses. RDF (Resource Description Framework) is a metadata framework providing interoperability between Web applications exchanging machine-understandable information.

**Full Text:***Copyright Online, Incorporated Aug 1998*

Most searches on popular Internet search engines yield hundreds or thousands of results, few of which correspond to what the query was intended to find. Why can't the engines do a better job of giving us what we want? Because HTML was defined as a presentation language, not as a way to structure document information. Although structure was added to support such objects as tables, the few structures available only make us want more, such as forms support. And no standard can satisfy every need. Thus XML, a standard for creating markup languages (like SGML, but Web-enhanced), was a response both to the explosion of Internet content and the need to create tags as needed.

Yet even with the potential for greater flexibility and specificity XML offers, a basic problem remains: How can I identify the attributes of an object (such as a text string), to say whether that string represents a word, a number, or a date? Without such information, XML-coded sites will be structured, linked, and styled without providing the necessary "metadata" tags to communicate what their information means.

## ENTER THE META MANAGERS

ICE, or Information Content and Exchange, is a protocol aiming to develop a consistent vocabulary for describing and managing the exchange of content and electronic assets between Web businesses. Web sites based on ICE will facilitate electronic commerce, including Web superstores. For example, a Web travel site could license other Web-based restaurant and hotel guides, vacation club materials, and similar travel-related content. Companies to

watch that will facilitate this new generation of business Web commerce include [Vignette Corporation](#) with its Story Server product and Firefly Network, a Cambridge, Massachusetts-based private company facilitating personalized and secure Web business relationships.

ICE will also facilitate syndicating content, helping publishers increase sales by making content licensing easier. Moreover, content buyers could add their own value and redistribute the content. For example, restaurant reviews could be distributed to newspapers as well as CD/Web media and even to cable outlets. Each could add its own value to that content via formatting, enriching links, integrating with other content, and so on.

Even more abstract is the Resource Description Framework, or RDF. RDF is a metadata framework providing interoperability between Web applications exchanging machine-understandable information. In fact, RDF can be applied to any resource that can be named by a Uniform Resource Identifier (URI—a more generalized way of pointing to Web resources than a URL). RDF creates a framework for creating, manipulating, and searching information. Over time, RDF could transform the Web into a coherent **digital library**. However, RDF itself is only a framework, and specific implementations (called vocabularies) must be developed to fulfill its promise. RDF is championed by [Microsoft](#) (RDF builds on [Microsoft's XML-Data proposal](#)), Netscape (**building** on Netscape's Meta Content Framework), content-providers like CNN and ABC News, and search systems suppliers like Alta Vista and [Yahoo](#).

One leading group building on RDF is called the Dublin Core, an international effort to make much of the Web into a modern-day equivalent of the ancient, grand Library of Alexandria. Dublin Core ([http://purl.org/metadata/dublin\\_core](http://purl.org/metadata/dublin_core)) is a set of 15 metadata elements designed to facilitate discovery of electronic resources. Although this clearly emphasizes Web sources, it is not intended only for Web content. Moreover, the type of resources being described are not restricted to text; they may include multimedia types such as pictures and sound. This effort has attracted the attention and talent of international resource description groups, such as museums and libraries.

Dublin Core's 15 metadata elements, each optional and repeatable, are grouped into three categories: content, intellectual property, and instantiation. The seven content tags include the familiar Title, Subject, Description, and the like. The four intellectual property tags are Creator, Publisher, Contributor, and Rights. Lastly, instantiation includes Date, Type, Format, and Identifier.

The list of international projects using Dublin Core is long and growing. Samples include Europe's Euler (integrated bibliographic databases, academic journals, and mathematical Internet sources), Germany's Subject Area Information for Earth (Earth Sciences information on Internet servers, CD-ROMs, and reference books), and Florida International University Digital Library (cataloging images, sound, and video for all subjects in the university's teaching and research portfolios).

## HOW META CAN GET BETTER

What impact will metadata initiatives such as RDF have on content developers and search systems? By guaranteeing metadata to be interoperable and able to be processed by software automatically, these metadata standards will provide a uniform way to search for information. Clearly this will enhance information searches by providing more precise results than even the best queries to a Web search system can now provide. And if you provide content, for the Web or elsewhere, you will want to be sure your content is accessible to as many users looking for it as possible.

### [Author note]

Robert J. Boeri ([bboeri@world.std.com](mailto:bboeri@world.std.com)) and Martin Hensel ([mhensel@hensel.com](mailto:mhensel@hensel.com)) are cocolumnists for INFORMATION INSIDER. Boeri is an Information Systems Publishing Consultant at a Boston-area loss prevention and control service company. Hensel is founder of Martin Hensel Corporation, a Newton, Massachusetts-based consulting firm that builds SGML-based editorial and production systems for publishers, corporations, interactive services, and composers. Comments? Email us at [letters@onlineinc.com](mailto:letters@onlineinc.com), or check the masthead for other ways to contact us.


---

Reproduced with permission of the copyright owner. Further reproduction or distribution is prohibited without permission.

ProQuest®

[Return to NPL Web Page](#)[Text Version](#)

English

[?Help](#) Collections Search  
Methods Topic  
Finder Browse  
Lists Results &  
Marked List Search  
Guide

Searching collections: All Collections

Article Display

[Email Article](#)

⊕ Article 36 of 42 ⊕

[Publisher Info.](#)[Print Article](#)☐ Mark article

Article format: Text+Graphics

[Save Link](#)

Saves this document as a Durable Link under "Results-Marked List"

## Building the digital library

*Federal Computer Week*; Falls Church; Aug 24, 1998; [Elana Varon](#);**Volume:** 12**Issue:** 29**Start Page:** 1, 44**ISSN:** 0893052X**Subject Terms:** [World Wide Web](#)[Libraries](#)[Information retrieval](#)[Research & development](#)[R&D](#)[Federal funding](#)**Geographic Names:** United States

US

**Companies:** [National Science Foundation](#)**Abstract:**

*While the World Wide Web has given federal agencies unprecedented access to reference data, it has also created a new challenge: finding the tools that make sense of that information. Agencies ultimately want the capability to pull text, audio, video and images from virtual shelves anywhere in the world as easily as one picks books from a library shelf today. Government-backed projects at universities and within federal agencies are attempting to develop digital libraries.*

**Full Text:***Copyright Federal Computer Week Aug 24, 1998***[Headnote]**

\* Curators of online collections are looking for ways to mine the information explosion

While the World Wide Web has given federal agencies unprecedented access to historical, scientific and reference data, it has also created a new challenge: finding the tools that make sense of that information. Whether they are homing in on key intelligence data or supplying educational materials for teachers, agencies ultimately want the capability to pull text, audio, video and images from virtual shelves anywhere in the world as easily as one picks books from a library shelf today

The Web "is what most people consider their digital libraries," said Michael Lesk, author of the recent book *Practical Digital Libraries* and director of the National Science Foundation division that manages the interagency Digital Libraries Initiative (DLI). Now curators of online collections are looking for ways to mine this information explosion.

"One of the classic functions of a library as an organization is they collect and acquire, they organize and make accessible, and they preserve," said Clifford Lynch, executive director of the Coalition for Networked Information, an interest group that represents universities and research libraries on technology issues. "If you look at many Web sites, they're about publishing information but not really so concerned about long-term retention and organization."

Librarians in the Map and Imagery Laboratory at the University of California at Santa Barbara (UCSB) began

"looking for ways to offer greater accessibility to their holdings of more than 5 million topographical maps, satellite data, photographs and other resources more than a decade ago. Back then, "no one even knew what we were talking about," said Larry Carver, the lab's director, when the library's staff said they needed a way to catalog, search and distribute their data online. "The technology was not there yet."

With help from federal grants, the university last month took the first step toward opening its holdings via the Internet. UCSB's Alexandria Digital Library (ADL) became the first link in an effort throughout California to provide electronic research materials on the Web first within the state university system and, after a year of testing, to the public, including agencies, universities and private companies around the world.

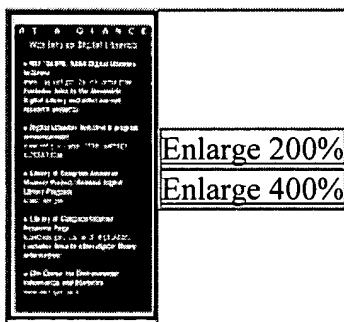
Other government-backed projects, at universities and within federal agencies, are chasing related ends. Through the DLI, during the past four years, NSF, the Defense Advanced Research Projects Agency and NASA jointly have poured \$24.4 million into ADL and five other academic projects that aim to make digital libraries as user-friendly as their physical counterparts.

Any organized collection of electronic documents that is set up "for human beings to use" can be a digital library, according to NSF's Lesk, and some technologies to support those collections are well-established. Databases, software for capturing images, tools for searching and retrieving text, and CD-ROMs and networks for distributing data are widely used by federal agencies today.

But technologies for tapping sound and video files, or for parsing data in multiple formats, have begun to emerge only recently. Developing search tools, including better ways to tag and index data, has been a major focus of digital libraries research. "I need to have ways to describe what I'm looking for," said Nand Lal, who oversees a NASA research project, called Digital Library Technology, that is separate from the joint effort with NSF. That means making satellite imagery and other space data currently organized for in-house use "intelligible" on the Web for scientists and the general public. "It's the whole idea of universal access in the sense of being able to deliver things that are of interest to the user, not necessarily to the producer, using facilities and language and terminology that are adapted to the user," Lal said.

"Search and retrieval is no longer about text," said Mark Demers, director of marketing and corporate communications with Excalibur Technologies, which this week is releasing software for searching video archives. Demers thinks the software could be used by federal agencies to set up libraries of training materials, surveillance tapes and historical records. "It's about all assets everywhere, and it's even about metadata," which is the text or software codes used to index online materials. Robust, easy-to-use search tools are "an enabling technology that is almost at the core of a digital library system," Demers said.

"An ideal goal for these technologies would be to make them disappear," said Stephen Griffin, the DLI program director, who is preparing to award a new round of grants, totaling \$40 million to \$50 million, beginning this fall. "If these technologies were invisible to the user and the user could work directly with the [information]," then the user could more easily create, and learn from, new virtual environments. Meanwhile, agencies are **building** basic **digital libraries** with software already on the market. Brand Niemann, **digital** librarian with the Environmental Protection Agency's new Center for Environmental Information and Statistics (CEIS), recently collected more than five dozen links for the center's Web site - links that enable users to access reports and data about local, national and international environmental conditions. Users can search these links, or only a portion of them, using the Topic search engine from Verity Inc.



One feature that makes the site a digital library, Niemann said, is that it offers visitors a single query form to search many sources, even documents hosted on other agencies' Web sites. Niemann also has helped the U.S. Geological Survey develop a "Web-connected CD-ROM" that gives users a set of documents they can use offline but that contains links to a Web site where users can obtain updates. The USGS application is based on digital publishing software from Folio Corp. that allows access to documents in different formats through a common interface.

Funding: The Biggest Barrier Niemann said funding, more than technology, has limited what CEIS has been able

"to include in its online collection." The content is endless. You'll never have it [all], so just like building the Web, you've got to get so many people involved."

According to NSF's Lesk, economic constraints, together with legal obstacles faced by agencies that want to distribute copyrighted data, form the main barriers to setting up digital libraries.

"Government libraries hold lots of materials to which they don't have the intellectual-property rights," said Bob Zich, director of electronic programs with the Library of Congress' National Digital Library program. Legislation pending in Congress aims to set rules governing copyrights online, but librarians and researchers, including an LOC official, have testified that these proposals could hamper public access to electronic materials in library collections.

Accessibility is another hurdle that agencies face. LOC started distributing copies of historic photographs and documents on disc eight years ago and now makes about 500,000 images, maps, film clips, audio files and texts available through its American Memory Web site. The goal of the \$60 million project, which is funded mainly through private donations, is to provide broader public access to historic and cultural artifacts from library collections around the country -- collections that otherwise would be accessible only by visiting the places where they are stored.

But although software such as RealAudio and QuickTime theoretically put sound and video clips within reach of anyone with Internet access, Zich noted that unless someone has a high-speed link, he might not have the time or the patience to download these files.

"We are waiting for when millions of people will have real wideband access," Zich said. "We have some films of the [1906] San Francisco earthquake that are loom."

Digital librarians also want more robust tools behind their collections' home pages. One such tool, under development by the NSF-backed San Diego Supercomputing Center and IBM Corp., offers a method for retrieving documents from different storage platforms. Much scientific data is stored as flat files, said Chaitanya Baru, senior principal scientist for enabling technologies at SDSC. "We haven't seen too many people trying to address this issue of trying to get to data on heterogeneous storage devices," he added. The project, part of a test bed for a paperless system to apply for patents, integrates IBM's High Performance Storage System, which is used to manage large data files for supercomputing applications, with a relational database. To find files, users query a database that holds a metadata directory, and a middleware application that the team has developed communicates the query to remote storage systems.

There is no single technology that appears to be driving the deployment of digital libraries today. But researchers and industry experts agree on the ultimate goal: that people must be able to find what they seek.

"Things can be put in formats that are lost forever," said Eugene Miya, a NASA electronics engineer who is reviewing DLI grant proposals. "If you don't have the protocols and formats [to retrieve it], your data is just about useless"

---

Reproduced with permission of the copyright owner. Further reproduction or distribution is prohibited without permission.



# **A T A G L A N C E**

## **Web Info on Digital Libraries**

- **NSF/DARPA/NASA Digital Libraries Initiative**

[www.cise.nsf.gov/lis/dli\\_home.html](http://www.cise.nsf.gov/lis/dli_home.html)  
(includes links to the Alexandria Digital Library and other current research projects)

- **Digital Libraries Initiative II program announcement**

[www.nsf.gov/pubs/1998/nsf9863/nsf9863.htm](http://www.nsf.gov/pubs/1998/nsf9863/nsf9863.htm)

- **Library of Congress American Memory Project/National Digital Library Program**

[lcweb.loc.gov](http://lcweb.loc.gov)

- **Library of Congress Internet Resource Page**

[lcweb.loc.gov/loc/ndlf/digital.html](http://lcweb.loc.gov/loc/ndlf/digital.html)  
(includes links to other digital library information)

- **EPA Center for Environmental Information and Statistics**

[www.epa.gov/ceis](http://www.epa.gov/ceis)

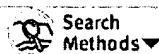
ProQuest®

[Return to NPL Web Page](#)[Text Version](#)

English

[?Help](#)

Collections

Search  
MethodsTopic  
FinderBrowse  
ListsResults &  
Marked ListSearch  
Guide

Searching collections: All Collections

Article Display

[Email Article](#)

⊕ Article 160 of 162 ⊕

[Publisher Info.](#)[Print Article](#)☐ Mark article

Article format: Full Text

[Save Link](#)

Saves this document as a Durable Link under "Results-Marked List"

# San Francisco Chronicle

## Archiving the Internet / Brewster Kahle makes digital snapshots of Web

San Francisco Chronicle; San Francisco, Calif.; May 7, 1998; [Said, Carolyn](#);

NAICS:7375 NAICS:7375

Start Page: B.3

Subject Terms: [Search engines](#)  
[Web sites](#)  
[Archives & records](#)  
[Entrepreneurs](#)  
[Corporate profiles](#)

Geographic Names: San Francisco

CA  
US  
Pacific

Personal Names: [Kahle, Brewster](#)

Companies: [Internet Archive](#) NAICS:7375  
[Alexa Internet](#) NAICS:7375

**Abstract:**

*His nonprofit Internet Archive serves as a historical record of cyberspace. His for-profit company, Alexa Internet, uses the archive as part of an innovative search tool that lets users call up "out-of-print" Web pages.*

*From a 100-year-old, red-roofed office in the Presidio, Alexa's 32 employees send out computer programs that crawl the Internet to find and download Web pages. It takes about two months to capture the entire Web -- currently some 300 million pages.*

*Along with the actual pages, the programs retrieve and store "metadata" --information about each site, such as how many people visited it, where on the Web they went next and what other pages are linked to it.*

**Full Text:**

Copyright Chronicle Publishing Company May 7, 1998

Brewster Kahle is creating the Internet equivalent of the Library of Congress.

The 37-year-old programmer and entrepreneur has been capturing and archiving every public Web page since 1996.

His nonprofit Internet Archive serves as a historical record of cyberspace. His for-profit company, Alexa Internet, uses the archive as part of an innovative search tool that lets users call up "out-of-print" Web pages.

2 MONTHS TO CAPTURE

From a 100-year-old, red-roofed office in the Presidio, Alexa's 32 employees send out computer programs that crawl the Internet to find and download Web pages. It takes about two months to capture the entire Web -- currently some 300 million pages.

Along with the actual pages, the programs retrieve and store "metadata" --information about each site, such as how many people visited it, where on the Web they went next and what other pages are linked to it.

The Web pages are stored digitally on a "jukebox" tape drive the size of two soda machines. It contains 10 terabytes of data -- as much information as one-half the entire Library of Congress.

Like that institution, the Internet Archive doesn't exclude information because it's trivial, dull or just plain weird.

## A VIRTUAL LIBRARY

"Of course, we've got more pictures of Cindy Crawford than the Library of Congress does," said Kahle. But to create an accurate portrayal of our life and times, it's necessary "to capture all the dreck you could ever want."

Having created a virtual library, the next step was to make a better card catalog. So Kahle and partner Bruce Gilliat started Alexa, named after the ancient Library of Alexandria.

Alexa's search engine uses the Archive's metadata to help users find information based on the trails of other Internet surfers.

The search engine, available for free at [www.alexas.com](http://www.alexas.com), is a toolbar that sits along the bottom of a Web browser. It looks at the site a user is currently viewing and suggests other pages by analyzing where previous visitors to that site went next.

## OLD SITES TO VIEW

What separates Alexa from other search engines is that it lets users view sites that have been removed from the Web.

When they encounter the message "404 Document Not Found" users can click on the Alexa toolbar to fetch the out of print Web page from the Internet Archive.

Alexa is supported by advertising, but even the ads relate to users' interests. A visitor to the [Amazon.com](http://Amazon.com) Web site might see a [Barnes & Noble](#) ad.

"Clearly we need better tools for exploring the Web," says Peter Lyman, head librarian for the University of California at Berkeley and an Internet Archive board member. "Alexa is trying to help us find our way out of the forest by looking for trails where previous people have gone. It's the most promising idea about how we'll search the Internet in the future."

## GRANDER PLANS

Available since September, Alexa already has 100,000 users but Kahle has grander plans for it.

"Our goal is to make this part of the infrastructure of the Internet," he said.

One surefire way to achieve that status would be to sell Alexa to a browser company, a search engine company or a major Internet service provider -- any of which might be a possibility, Kahle said.

Browser and search firms are snapping up technology that improves Web navigation. Search company Lycos last week spent \$39.75 million for WiseWire, which automatically organizes Internet content into directories and categories. Last month [Microsoft](#) shelled out a reported \$40 million for Firefly, which recommends content to Web surfers based on profiles they submit.

Kahle already has a track record of creating next-step Internet technology. In the early 1990s, he developed the Wide Area Information Server (WAIS), the first system for publishing quantities of data in a searchable form on the Internet.

## IMPRESSIVE BACKGROUND

The [New York Times](#), Wall Street Journal and [Encyclopaedia Britannica](#) were among its customers. Kahle later sold WAIS to [America Online](#) for \$15 million in 1995.

Besides an impressive programming background, which includes a degree from the Massachusetts Institute of Technology and a stint designing supercomputers at Thinking Machines Corp., Kahle has an abiding interest in traditional media.

His hobby is letterpress printing. Painstakingly aligning individual lead letters by hand to make cards and documents is a far cry from computer automation, "but that's the charm," he said.

His wife, Mary Austin, is the founder and curator of the San Francisco Center for the Book, which runs programs and classes to encourage "all arts of the visible word."

#### TYPE DESIGNER'S LEGACY

They named their 3 1/2-year-old son Caslon after an 18th century type designer. Their 9-month-old son Logan has a family name.

"When the printing press came about, it fostered thousands of tiny presses all over the globe, allowing people in small towns to publish and distribute information. That's what we're finding here on the Web," he said.

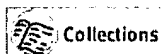
"As we move human knowledge from paper to computers, people are getting access to huge amounts of information more easily. But to help organize the Web we have to track what's on it and what's going on over time."

Textual Illustration:

PHOTO; Caption: In their office at S.F.'s Presidio, Alexa co-founders Brewster Kahle (left) and Bruce Gilliat stood amid their machinery. / BY JERRY TELFER/THE CHRONICLE

---

Reproduced with permission of the copyright owner. Further reproduction or distribution is prohibited without permission.



Collections



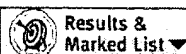
Search Methods



Topic Finder



Browse Lists



Results &amp; Marked List



Search Guide

Searching collections: All Collections

Article Display

[Email Article](#)

Article 26 of 31

[Publisher Info.](#)[Print Article](#)☐ Mark articleArticle format: [Text+Graphics](#)[Save Link](#) Saves this document as a Durable Link under "Results-Marked List"

## The dynamics of classification systems as boundary objects for cooperation in the electronic library

*Library Trends*; Urbana; Fall 1998; [Hanne Albrechtsen](#); [Elin K Jacob](#);

Volume: 47

Issue: 2

Start Page: 293-312

ISSN: 00242594

Subject Terms: [Library science](#)

[Classification](#)

[Information technology](#)

### Abstract:

*The notion of the classification scheme as a transitional element or "boundary object" (Star, 1989) offers an alternative to the more traditional approach that views classification as an organizational structure imposed upon a body of knowledge to facilitate access within a universal and frequently static framework.*

### Full Text:

*Copyright University of Illinois at Urbana, Graduate School of Library & Information Science, Publications Office Fall 1998*

### [Headnote]

ABSTRACT

### [Headnote]

THE NOTION OF THE CLASSIFICATION SCHEME as a transitional element or "boundary object" (Star, 1989) offers an alternative to the more traditional approach that views classification as an organizational structure imposed upon a body of knowledge to facilitate access within a universal and frequently static framework. Recognition of the underlying relationship between user access and the collective knowledge structures that are the basis for knowledge production indicates the dynamic role of classification in supporting coherence and articulation across heterogeneous contexts. To this end, it is argued that the library should be an active participant in the production of knowledge, and that this role can be effected by the development of classificatory structures that can support the needs of a diverse information ecology consisting of a complex web of interacting agents, users, and technologies. Within such an information ecology, a classificatory structure cannot follow a one-size-fits-all paradigm but must evolve in cooperative interaction between librarians and their user groups.

### INTRODUCTION

A bibliographic classification system is intended to provide both an overall structure for a document collection and a set of concepts that will guide the information searcher into the knowledge domains encompassed by the collection. Traditionally, classification research has approached these objectives by developing schemes based on a one-size-fits-all-searchers paradigm-i.e., We have created a standard system, because, deep down, all users are the same. Such classificatory tools often fail to fulfill their function of supporting the searcher's access to, and navigation through, the domain structure. In most databases, including catalogs on the Web, the searcher may find it difficult to comprehend the organizational structure that has been imposed upon the materials. This is not

due simply to the often exotic relations of a scheme or to the surface characteristics of the classificatory data. Rather, the problem is often a product of a lack of match between the structure imposed upon the retrieval system by the classification scheme and the user's individual knowledge structures and search strategies.

Classification research has responded to this problem by collecting the terminology of individual users and compiling the results to generate larger, broader, and, it is hoped, more successful sets of access points for users-i.e., If we design an end-user thesaurus, that should do the trick. In his recent book on information seeking and subject representation, Hjørland (1997) argues that such endeavors to compile end-user vocabularies are generally conducted without recourse to an underlying theory of knowledge. Because failure of the classificatory structure to support user access is generally interpreted as a mechanical question of matching between different individual knowledge structures-i.e., among those of the searcher, the author, and the librarian as mediator (compare, for example, Ingwersen, 1992)-the underlying relationship between user access and the collective knowledge structures that are the basis for knowledge production has not been widely recognized.

From the perspective of the sociology of science, Star (1989) has argued that the Turing test, which is intended to measure the degree to which an expert system is able to perform as a human expert in its interaction with individual users, should be replaced by a "Durkheim test," where the system is evaluated on its ability to support the goals of a specific community of users. Star points out that scientific work is not all one piece but is distributed and heterogeneous, with differing viewpoints emerging only to be reconciled within the existing knowledge base. In her view, information systems should not be designed simply to represent consensus but to accommodate the dissent that can be expected to appear among the various communities participating in their use. To this end, she brings forward the concept of boundary objects as a method for resolving problems of heterogeneity in knowledge production and use or, in terms of library and information science (LIS), problems of variation or inconsistency in the representations by information producers, information mediators, and information users.

In this article, we will investigate how classificatory structures can act as transitional elements or boundary objects (Star, 1989) to support coherence and articulation in the heterogeneous and sometimes distributed contexts where knowledge is produced and mediated. In particular, we will review, within the context of the library, two perspectives put forward by Hjørland (1997) and by Star (1989) that analyze information systems as dynamic social constructs. We will build an analogy between a scientific enterprise and the library as an active participant in the general production of knowledge and use this analogy to develop a view of modern classification research that engages the library directly in the development of classificatory structures that can accommodate information searching by heterogeneous user groups. Following Nardi and O'Day (1996), we regard the library as a diverse information ecology, comprising a complex web of interacting human agents, users, and technologies. And we will argue that, within such an information ecology, a classificatory structure cannot follow a one-size-fits-all paradigm but must evolve in cooperative interaction between librarians (and other information intermediaries) and their user groups. In this context, we draw on examples of information systems in Danish public libraries-i.e., the Book House (Pejtersen, 1980) and Database 2001 (Albrechtsen, 1997).

## CLASSIFICATION SYSTEMS: FROM RATIONALISM AND EMPIRICISM TO SOCIAL CONSTRUCTIVISM

Hjørland (1997) argues for a philosophical and sociological orientation for classification research. In his view, the problem of the searcher's uncertainty is a function of relative task uncertainty in the user's problem domain. Because information searching takes place within a particular social framework-e.g., an academic discipline-task uncertainty in searching is often the result of the relative task uncertainty within the discipline itself. Albrechtsen and Hjørland (1994) have earlier shown how such task uncertainty within knowledge domains may be a function of various social factors involved in the production of knowledge, such as the degree of interdisciplinarity or maturity within a domain. Such uncertainties will not only be manifest in the searchers' difficulty in formulating queries for IR-systems but will also be inscribed in the relative plasticity and variety of the concepts and terminology applied within the domains.

Classification research has too often neglected such broader social backgrounds that inform information searching and knowledge organization and has relied, more or less implicitly, on either a one-size-fits-all paradigm (rationalism) or on the accumulation of data about user behavior (empiricism). While the rationalist approach argues that we just need to get everyone to understand this, the empiricist counters that we just need to get more data about users and proceeds to collect more or less meaningful sets of "facts" on the individual user's relative success measured as the number of "hits" resulting from a series of search queries.

Figure 1 divides the different approaches to classification research and practice into two broad epistemological categories: Rationalism/ Empiricism on the one side and Historicism/Social Constructivism on the other. Both rationalism and empiricism are based on assumptions regarding the nature of truth and the objectivity of knowledge. From the empiricist approach, knowledge is reduced to sensory observations or facts. In classification research, empiricism is the prevalent epistemology in bottom-up thesaurus construction based either on user warrant or on terminology warrant, particularly when the process lacks grounding in a theory of knowledge. In contrast, rationalism strives to reduce knowledge to an all-embracing structure of concepts that is intended to be universally comprehensive. It is, for example, the epistemological foundation for Ranganathan's notion of universal

facets. Rationalism is also closely related to more sociopolitical actions undertaken by a particular agency or from a specific disciplinary viewpoint-i.e., actions which are intended to impose one view of knowledge on all research and practice within that domain. In a paper discussing the role of dialogue in the development of classificatory structures, Jacob and Albrechtsen (1997) have shown how the American Psychiatric Association's construction of DSM-IV (American Psychiatric Association, 1994), the international classification for mental disorders, used dialogue to create a device for marginalizing and eliminating the viewpoints of competing professions such as psychology (see also Kirk & Kutchins, 1992). In short, both empiricist and rationalist approaches to classification are primarily looking for invariant structures that can be imposed on encyclopedic knowledge (universalist approaches) or data compiled from local observations (e.g., grounded theory approaches).

In contrast to these more formalized structure-seeking approaches to classification, social constructivism, or historicism, offers a view of knowledge as a product of historical, cultural, and social factors, where the fundamental divisions and the fundamental concepts are products of the divisions of scientific/cultural/social labor in knowledge domains. According to a social constructivist epistemology, the concepts and the structures are inseparable in a classification system, and hence the schemes must reflect the development, variety, plasticity, and use of both within a particular knowledge domain. This implies that scheme designers are not primarily looking for ways to impose one single structure on knowledge, including one set of all-embracing facets. Rather, the designers should operate as "epistemic engineers," attempting to articulate and represent the dynamics of knowledge in such a way that the searcher can proceed from the topic of his initial query to other related perspectives on the same topic or to related materials within the same knowledge domain. In this manner, epistemic engineering of classificatory schemes can provide for multidimensional classification schemes where the concepts are represented in a variety of different conceptual structures, functioning to articulate the multiple discourses performed in different domains. In the role of epistemic engineer, then, the scheme designer operates as an active participant in the process of knowledge production and mediation.

Such involvement on the part of the classificationist is particularly evident in areas of interdisciplinary research that engage participation from many different professions. The HIV/AIDS vocabulary, developed by Huber and Gillaspay (1996), provides an illustrative example of such involvement on the part of the scheme designers. This system, which was not intended as a classification per se but as a mediating vocabulary, was developed to support dialogue between the different communities involved with the HIV/AIDS epidemic, including clinical and medical researchers, practitioners of alternative medicine, nutritionists, psychotherapists and other professionals, as well as those individuals who are either living with the disorder themselves or are caring for someone who has contracted the disease. The HIV/AIDS vocabulary is built on a theory of knowledge generation that explicitly eschews the standard life cycle for knowledge production in medicine-a knowledge cycle that proceeds in a top-down fashion from theory developed at universities and other research institutions, to applied clinical research, to daily clinical application. Rather, according to the epistemological view driving the HIV/AIDS vocabulary, research in lived experience must necessarily feed into basic clinical research. Accordingly, this scheme was not developed solely as a tool for retrieval of information in the database of the local community, but as a tool for facilitating communication both within and across diverse interest groups, from the so-called layman to the cloistered scientist. In its role as communicative facilitator, the scheme is also hospitable to adaptations and extensions as an indexing language in local contexts. For instance, specific drug names are not articulated in the scheme but are left to local instantiations of the indexing language. In Star's (1989) terms, the HIV/ AIDS scheme serves as a boundary object precisely because it supports cooperation and common understandings among the various interest groups touched by this epidemic.

## CLASSIFICATION AND BOUNDARY OBJECTS

The notion of "boundary objects" was developed by Star (1989) as a structure for coordinating distributed work, such as may occur with a scientific enterprise that not only involves heterogeneous actors, elements, and goals but also incorporates different research methods, values, and languages. From her field work with scientific communities, Star has found that scientists are able to cooperate without consensus or shared goals. They can work together successfully because they create objects that function in the same way as a blackboard in a distributed artificial intelligence system:

I call these boundary objects, and they are a major method of solving heterogeneous problems. Boundary objects are objects that are both plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites. They are weakly structured in common use, and become strongly structured in individual-site use.

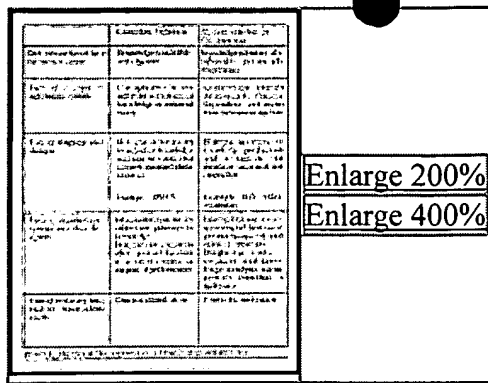


Figure 1.

Like a blackboard, a boundary object "sits in the middle" of a group of actors with divergent viewpoints. Crucially, however, there are different types of boundary objects depending on the characteristics of the heterogeneous information being joined to create them. (Star, 1989, pp. 46-47. Emphasis in original)

Accordingly, Star (1989; Star & Griesemer, 1989) has identified different types of boundary objects in her various case studies, including: repositories-databases, libraries, or museums; ideal types or platonic objects-diagrams, atlases, or abstract concepts such as, for example, the concept of "species" used by both the creators of a zoological museum and other interested parties involved in its construction;

coincident boundaries-common objects with the same boundaries but having different internal contents, such as maps of a geographical area that cover the same terrain but are outlined according to different knowledge interests such as, for example, the life zones identified by biologists contrasted with the trails and collection sites identified by museum conservationists;

standardized forms-forms created as methods of common communication across distributed work groups such as, for example, the forms completed during field work or the cataloging formats used for cooperation and networking between libraries where the content fields may or may not be part of each repository's database.

Unlike the model of the ideal universal computing machine whose goal, as proposed by Turing, is to emulate individual human mental capacities in all domains, boundary objects are advanced as an ecological concept-i.e., a concept that respects local contingencies and the viewpoints of different knowledge interests. In a case study on the formation of Berkeley's Museum of Vertebrate Zoology (Star & Griesemer, 1989), a classificatory structure of the species and subspecies of mammals and birds in California constituted an important boundary object. Thus the scientific classification scheme served as a shared conceptual structure and provided a shared vocabulary that facilitated communicative exchanges and cooperation across the different social and intellectual worlds represented by the scientists and the groups of amateurs who were involved in building the museum's collection.

Although they approach the problem of classificatory structures and knowledge access from two different angles, Star's exposition of the communicative and integrative functions of classificatory structures in the general knowledge production of the sciences is closely related to Hjørland's (1997) discussion of the epistemological positions adopted in classification research and his argument for following a more pragmatic philosophy of classification. Star builds on case studies and theoretical work in scientific communication and knowledge production, while Hjørland builds on case studies and theoretical work in the area of information searching for knowledge production. Both argue that classifications always serve pragmatic purposes in the same way that science serves human action. According to Hjørland's theory, scientific classifications reflect a highly abstract and generalized method of knowledge organization, in contrast to classifications with more local contingencies, such as categorizing fruit and vegetables in a supermarket or the amateur horticulturist's categorization of plants by use or cultural preferences. Such variations in taxonomic structure could be argued to reflect different levels of ambition among the interested parties and thus to function as boundary objects, created and negotiated by different social worlds, with the scientific structure functioning as a more specific taxonomy dictated by the needs of the scientific community itself. However, with respect to its specific role within the praxis of a formal disciplinary community, the scientific taxonomy is just as concrete as the pragmatic systems of classification that reflect local contingencies. Indeed, when viewed from a broader sociological perspective, these latter systems may actually be interpreted as more abstract or generalized.

## THE ROLE OF CLASSIFICATIONS IN DIVERSE INFORMATION ECOLOGIES

American anthropologists Nardi and O'Day (1996) have introduced the concept of "diverse information ecology" to describe the sociotechnical network of heterogeneous materials, people, and practices that constitutes a modern library:



What we learned in the library suggests the possibility of a socio-technical synthesis, an opportunity to design an information ecology that integrates and interconnects clients, human agents and software agents in intelligent ways congenial to extending information access to, potentially, all of humanity. As we design the global information infrastructure, the ultimate goal should be to design an ecology, not to design technology. (p. 83)

Because information ecologies are situated within human practice, they are dynamic and constantly changing. An information ecology cannot be controlled by any one single agency but evolves through the collaboration of heterogeneous socio-technical networks, whose elements strive constantly to achieve coherence and wholeness. The notion of an information ecology also implies a collective view of information systems as striving to meet heterogeneous community goals rather than the goals of a single agency or individual. In their study of two research libraries in software companies in the United States, Nardi and O'Day (1996) explored how the work practices and expertise of librarians can serve as a model for the design of computerized information services. They found that librarians are exemplary agents who evince particular expertise not only in communicating with users but also in searching for information. These two skills are closely interrelated in that the librarian's search strategy tends to evolve in collaboration with the user's project. Nardi and O'Day propose to extend this working relationship between the librarian and the user to the collaborative design of information ecologies.

In an information ecology, a classification system should function as a boundary object, supporting coherence and a common identity across the different actors involved. In its role as boundary object, a classification would be weakly structured in common use, while remaining open to adaptation in individual communities. Across diverse information ecologies, classification schemes would function as discursive arenas or public domains for communication and production of knowledge by all communities involved. This approach to the development of classification schemes also implies that the task of constructing such a scheme would no longer be invisible work. This view of classification systems is in line with the concept of "coordination mechanisms" in distributed collaborative work, as put forward by Schmidt and Simone (1996). More importantly, the understanding and appreciation of classification schemes as boundary objects and discursive arenas, in cooperation with heterogeneous user groups and technology, engages the library as a facilitator of connections and ensures its continuing participation as an active contributor in the general process of knowledge production.

In the following discussion, we will illustrate how the role of classification systems is changing within the information system that is the library, shifting from reliance on a single standardized or mainstream view of order, where classification is the invisible precursor to the organization of a collection, toward the creation of more diverse information ecologies, where the development of a classification scheme takes place within an arena of discourse to create a shared order across heterogeneous social worlds.

### SOMETHING OLD, SOMETHING NEW, SOMETHING UNIVERSAL, SOMETHING LOCAL

As indicated in Figure 2, classification systems have served different pragmatic purposes in the history of libraries and information retrieval systems. In a recent European study of public libraries in the information society (Thorhauge & Segbert, 1997), it was demonstrated that public libraries have progressed through three distinct stages, evolving from manual paper-based services, via the automated library, to the current phenomenon of the electronic multimedia library. This progression should not be understood to imply that the current status of libraries has been driven entirely by technology. Rather, the electronic multimedia library must be understood from a more integrated socio-technical point of view, where the various actors, including librarians, computer suppliers, and researchers in computing and information science, constitute a heterogeneous network of agencies that bring certain technologies to the foreground while marginalizing others. In the recent development and use of communication technology, for example, there is a convergence of hitherto separate, even disparate, media and activities. This is apparent in the development and application of Web technology, which integrates textbased materials, graphic illustrations, and audio materials with interactive features such as online conferences and e-mail. It is characteristic of this development that the technology is not only plastic and customizable to almost any context of use, rather like a boundary object, but is constantly renegotiated and redeveloped through such use.

In the recent past, manual paper-based libraries focused on collection building. Intermediaries, or librarians, served both as collection builders and as agents controlling and interpreting the order of the libraries. Classification systems were frequently standardized in order to support interlibrary cooperation with the result that classification research was itself dominated by the development of universal schemes which could be adopted by central agencies to control the organization of knowledge across libraries. As a result of such standardization, classification became invisible work performed without regard to the needs of the local community of users. And, because maintenance and development of these classification schemes was often based on literary warrant, reflecting only those subjects represented in large national collections, they can be interpreted as imposing an implicitly empiricist view of knowledge. There was, then, at this stage in the library evolution, a mix of rationalist and empiricist epistemologies underlying classification research and development.

The role of the librarian as intermediary was challenged during the 1980s by the development of online retrieval systems and, in particular, by the introduction of online public access catalogs (OPACs) for end-user searching.

[illegible]

During the 1990s, the **library** has increasingly switched its service emphasis from **building** and guarding the collection or offering users access to the collection through the local OPAC to providing local access to global information resources available on the World Wide Web. This represents a shift from a closed to an open system. In some European public **libraries**, for example, traditionally introverted and bureaucratic organizations have migrated toward a project-oriented culture, where librarians and users cooperate on the development of new services, using the interactive affordances of Web technology and the Internet. In general, such projects have not involved the **library** schools in Europe, the traditional research communities in the **library** and information sciences. Close cooperation between **libraries** and the community of LIS researchers in Europe has yet to be manifested (Albrechtsen & Kajberg, 1997) . In the United States, communities of LIS researchers have come together in workshops and research projects related to the social informatics of what are called "**digital libraries**" but could equally well be termed "electronic **libraries**" (Bishop & Star, 1996). In this research area, major topics include how knowledge is structured in **digital libraries**, including cataloging and classification, and how digital libraries are used-i.e., how knowledge is produced, communicated, applied, and recycled in distributed social worlds. Research methods comprise ethnographic studies of communication and knowledge production in (digital) libraries as well as comprehensive sociological studies of professional classification schemes in medicine (Bowker & Star, 1994) and nursing (Bowker, 1996). Thus it seems apparent that classification research is gradually evincing a more sociological and historical orientation.

Ballerup Public Library is a medium-sized Danish library on the outskirts of Copenhagen. There is, in this library, a tradition of direct collaboration between the librarians and their users. Until recently, a majority of the librarians regarded themselves as cultural workers—as intermediaries between collection and user, very much in line with the traditional perspective described above for libraries in the manual stage. In 1995, the library started a new project called Database 2001. This project, which was evaluated by Albrechtsen (1997), involved the development of an enriched multimedia catalog on the Web. In addition to the evaluation researcher, the project group for Database 2001 included six librarians with different areas of expertise: several in the group were experienced intermediaries and online searchers, while others were specialists in catalog design and in the management of the library's technological resources. However, none of the librarians had experience with Web design or Internet browsing.

2/11/03 4:11 PM

browsing the menus was customized for both children and adults. The librarians arranged evaluation sessions with users who represented different user communities and their evaluations were very positive; users with different interests were able to use the icon-based interface for browsing in the database even though they had very different interests and different goals for searching.

In the database, documents were indexed using standard call numbers from the Danish variant of the Dewey Decimal Classification (DDC). Even though indexing by class number would take advantage of the hierarchical structure of DDC and thus would be potentially useful for browsing by users, the librarians knew from their practice as intermediaries that users found it very difficult to understand the standard classification. They experimented with a more pragmatic and much more weakly structured classification which could reflect the kinds of questions actually posed to library staff by the different user groups. For example, for subject browsing by children, they worked with the seven categories listed below and designed a unique icon to be used on the Web site:

1. computers;
2. astronomy, nature, animals, environment;
3. first love, star signs, being young today;
4. horses;
5. excitement, humor;
6. fantasy, science fiction; and
7. books that are easy to read.

From a semantic or disciplinary point of view, the separation of subjects like animals and horses would appear to be "incorrect" or "illogical." For the children, however, this classification worked very well. Category 2 (astronomy, nature, animals, environment) was intended for a broad group of interests, including fact literature, whereas category 4 (horses) was intended, in particular, for girls interested in novels about horses. There is, in Denmark, a special research tradition within children's librarianship, based on Wanting's (1984) research on how children ask questions in libraries, that advocates mediating literature according to the different user interests of children. Pejtersen (1994) has also studied children's use of libraries in Denmark and their communication with librarians. In her development of the Book House system in the 1980s, Pejtersen used a collaborative prototyping approach, engaging librarians, information scientists, and users in Danish public and school libraries, and subsequently designed a special interface of subject icons for browsing of the Book House system by children. Database 2001 took advantage of both of these research approaches to children's information searching.

The Book House (Pejtersen, 1994) is a retrieval system for fiction and is based on a general conceptual model that seeks to surround users with an adequate resource space within which to situate their own search spaces. The design involves multidimensional representations of different kinds of user needs, search strategies, and literary paradigms as well as authorial intentions. This multidimensional structure for subject access is intended to match the different levels of user interest. The system interface is constructed around the metaphor of a "house of books," guiding the users through the rooms of a library where they can browse the collection. Users can also switch between different search strategies, including analytical search in the multidimensional database structure, visualized as icons for each dimension, and browsing of subjects, visualized as icons in a picture gallery. The design of these icons involved classification experiments using both word association experiments and evaluations of suggested icons in Danish public libraries.

The icons for browsing subjects in the Book House and in Database 2001 serve similar functions-to provide the users with an overview of the subjects included in the databases. Because the Book House system builds on the central design metaphor of rooms in a library, it provides a single uniform interface. Database 2001, in contrast, is realized as a mixture of interfaces that include the Web layer of icons, designed by the librarians; a more or less standard search client offering conventional text-based searching; and a database structured according to a standard cataloging format that uses traditional call numbers to represent the subject content of the documents. While the Book House is a general system for fiction retrieval, which in its present form cannot be customized by individual libraries to support the idiosyncratic needs of specific user communities, Database 2001 is a localized experiment with system design and classification drawing upon a range of technologies that reflect the heterogeneity of tools used in today's libraries, from conventional customizable applications such as the closed systems of the database and the search client to the open systems supported by interactive Web technologies.

## COLLABORATIVE DEVELOPMENT AND THE AGENCY OF LIBRARIES

Both the Database 2001 project and the Book House system were realized using a collaborative approach among librarians, users, researchers, and technicians. In this way, users were involved in negotiating classificatory

structures and the design of subject icons in the interfaces of the two systems. Because the Book House was a new approach for interface and database design in the 1980s, it had to be developed technically from scratch. Database 2001, on the other hand, was able to take advantage both of the design ideas generated during development of the Book House system and of the possibilities for integrating modern Web capabilities within existing technology. The process of designing an interface adapted for local needs quite naturally involved local experiments with classification. In Database 2001, the graphic Web layer and its icons were intended to represent both the users' needs and the existing technology. Decisions regarding the subject icons, as well as those pertaining to the search client and the database, were determined as much by the users as by the demands of the Web technology itself. Thus the icons employed in the graphic interface constitute an integrated system of boundary objects that mediate among the library, the users, and the technology. In this way, Database 2001 exists as an open system in that it makes the library available not only to local users but to other users as well through the medium of the technology. Without the interface of icons, the system would have been technically open but conceptually closed.

Design of the Book House and Database 2001 involved heterogeneous human actors, elements, and goals, which are also found in Star's (1989) description of a scientific enterprise. Star draws upon the example of a scientific enterprise to put forward a more collective concept of design than the psychological approach generally employed for the design of AI systems. Traditionally, design of library systems is based on a consensus model, or a one-size-fits-all approach. Multidimensional classifications providing different views of concepts in IR systems are still the exception (Albrechtsen & Hjørland, 1994; Jacob, 1994). But in the Book House system and in Database 2001, classificatory structures can perform as boundary objects by accommodating both the heterogeneous information needs and the various search strategies of different user interests as well as different knowledge communities.

Figure 3 juxtaposes some important boundary objects developed in the Book House and Database 2001 with Star's typology in order to illustrate the analogy between boundary objects in a scientific enterprise and the creation of a library system. Obviously, this analogy between the library and a scientific enterprise, even when supported by parallel structures, does not establish that what goes on in a library is isomorphic to what goes on in a scientific enterprise. Hjørland (1997) has proposed a theory of classification at multiple levels, from specific classifications developed in accordance with local contingencies to those general classifications developed by the so-called "hard" sciences, such as biology and medicine. However, analysis of the role of dialogue in the creation of classificatory structures indicates that traditional classification schemes frequently function as control structures that forestall an interpretive approach to scheme design through the imposition of controlled vocabulary that limits the impact of dissonant viewpoints (Jacob & Albrechtsen, 1997). In this manner, current developers of classification systems do not function as epistemic engineers, creating a discursive arena or forum for multiple views of knowledge, but rather as engineers of one episteme or worldview seeking to control the flow of knowledge production within a given domain by systematically legitimizing a single universal classificatory scheme, thereby disenfranchising those researchers and practitioners who do not participate in the resulting structure.

In general, however, the library and its organizational structures must be viewed as important actors in the general process of knowledge production because their primary goal is to mediate between knowledge producers and users. This role is generally realized through the provision of information services to users and producers who are very often members of the same knowledge communities. Although the scenario sketched here is traditionally understood as a closed world-i.e., where librarians mediate between documents and users- it could equally well be described as an information ecology-i.e., as a practice that builds environments by bringing together heterogeneous materials and actors.

The librarians' practice of building information ecologies is based on both explicit and tacit knowledge. The explicit knowledge is typically based on principles and formalisms for presenting classificatory structures in the form of universal classifications or faceted thesauri. The tacit knowledge includes knowledge of the interests of their user communities, the users' levels of computer and information literacy, and preferred tactics for "mediated" versus "unmediated" information services. In mediated services, the librarians communicate with the users, either directly or by e-mail, and guide them to relevant information sources. In unmediated services, such as the Book House system or Database 2001, the users may search a card catalog, a database, or a contingent local classification scheme prior to browsing the conceptual space within a domain. Such "unmediated" services are, in fact, "silently" mediated by librarians or other information professionals who designed or customized a conceptual space for end-user searching. The librarian's service to the users has been "translated" or formalized through the classification scheme. It has been abstracted or "disembedded" from the work context of a human intermediary interacting with a user.

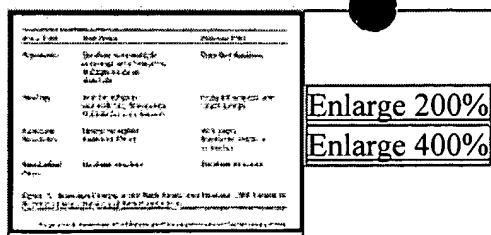


Figure 3.

Following Star and Strauss (1999), much of the mediating practice of librarianship may be considered "invisible work." Even though the librarian as human intermediary is visible within the traditional library setting, his or her work is frequently considered to be "background work" involving the identification and delivery of books or other materials in support of the "real thing"-i.e., the user's immediate work task or particular interest. When the work of the intermediary is abstracted from the work setting, this "invisible work" may become "visible" in that the task now falls to the user, but the dialogue between the user and the intermediary is effectively silenced. No longer is there a human intermediary to inform the user and ensure equality of services.

Gross and Borgmann (1995, cited in Bishop & Star, 1996) point out that: "Even home shopping requires informed consumers" (p. 904). When the librarian's mediation work is silenced in the high-tech home shopping environment of electronic libraries-when the task of the user is no longer supported by, or facilitated through, dialogue with the human intermediary-some users will not be informed but will be reduced to mere consumers of standardized information services. Obviously, then, the information ecology of the electronic library cannot be responsive to the needs of the individual user without achieving a balance between visible and invisible work. As Star and Strauss (1999) point out: "Making visible can incur invisibilities; obscuring may itself become a visible activity." In "unmediated" information services, cooperation between librarians and users in the design and maintenance of classificatory structures may be one method for achieving this balance between the visible and the invisible and for ensuring the evolution of an information ecology that is contingent upon the needs of an informed public.

## CONCLUSION

Classification systems and indexing languages have been constructed as organizational tools in order to provide structure to a body of knowledge, but they frequently have the effect of limiting or restructuring individual conceptual structures during a process of information searching (Tang & Solomon, 1998). Established approaches to classification research and development appear to suffer from a fear of touching the real thing-the social worlds constituting an information system and the collective conditions for knowledge production. However, in LIS and the sociology of science, new approaches to classification research are emerging, approaches that build on the idea of information systems as open and collaborative systems. A similar trend toward development of open systems has been identified in the public libraries in Europe which are evolving from manual, paper-based services to the electronic multimedia library. In the modern electronic library, classification is similarly transformed from a tool for establishment of order and control over the collection to a boundary object functioning to create cohesion across diverse information ecologies. The modern information ecology is a socio-technical network comprised of heterogeneous materials, people, and practices. Within this emerging network, the classification scheme constitutes a discursive arena facilitated by the library and functions as a boundary object for the various interests that exist among users and librarians. Such an information ecology is at the same time a situated network and an open system wherein the classification scheme supports coherence and articulation across the domains encompassed by the network both locally and globally.

The practice of classification is changing from invisible work carried out in centralized agencies to articulation work emerging within socio-technical networks. As the role of the library evolves from collection guardian to facilitator of connections, the role of classification is similarly transformed from control of collections to facilitation of communication, maintenance of coherence, and establishment of a shared conceptual context. From this perspective, then, the intelligent intermediaries of today are the human agents in diverse information ecologies who facilitate the process of knowledge production by collaborating with communities of users in the creation and use of boundary objects such as classification schemes.

## ACKNOWLEDGMENTS

We gratefully acknowledge the inspiration and comments by Susan Leigh Star for this article. We also want to acknowledge the Danish National Library Authority for support of the project Database 2001 in Ballerup Public Library. An earlier version of this article was presented and discussed at the ASIS SIG/CR workshop in Washington, DC in 1997.

## [Reference] REFERENCES

**[Reference]**

- Albrechtsen, H. (1997). Database 2001: Design af en multimedie Web katalog pa Ballerup Bibliotekerne. Copenhagen, Denmark: Royal School of Librarianship.
- Albrechtsen, H., & Hjørland, B. (1994). Understandings of language and cognition: Implications for classification research. In R. Fidel, B. H. Kwasnik, C. Beghtol, & P. J. Smith (Eds.), *Advances in classification* (Proceedings of the 8th ASIS SIG/CR Classification Research Workshop, Alexandria, VA, 16 October 1994) (vol. 5, pp. 1-16). Medford, NJ: Information Today.
- Albrechtsen, H., & Kajberg, L. (1997). Professional development: Education, training and continuing professional education. In J. Thorhauge & M. Segbert (Eds.), *Public libraries and the information society* (pp. 82-91). Luxembourg: Office for Official Publication of the European Communities.

**[Reference]**

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: APA.
- Bishop, A. P., & Star, S. L. (1996). Social informatics of digital library use and infrastructure. In M. E. Williams (Ed.), *Annual review of information science and technology* (vol. 31, pp. 301-401). Medford, NJ: Learned Information.
- Bowker, G. C. (1996). Lest we remember: Organizational forgetting and the production of knowledge. Retrieved September 1, 1998 from the World Wide Web: <http://alexia.lis.uiuc.edu/~bowker/pubs.htm>
- Bowker, G. C., & Star, S. L. (1994). Knowledge and infrastructure in international information management: Problems of classification and coding. In L. Bud-Frierman (Ed.), *Information acumen: The understanding and use of knowledge in modern business* (pp. xx). London, England: Routledge.
- Gross, R. A., & Borgman, C. L. (1995). *The incredible vanishing library*. American Librar

**[Reference]**

- ies, 26(9), 900-904.
- Hjørland, B. (1997). *Information seeking and subject representation: An activity-theoretical approach to information science*. Westport, CT: Greenwood Press.
- Huber, J. T., & Gillasp, M. L. (1996). Controlled vocabulary for HIV/AIDS: An evolving nosological record of a disease body of knowledge. In R. Green (Ed.), *Knowledge organization and change* (Proceedings of the fourth international ISKO conference) (pp. 293-301). Frankfurt/Main, Germany: Indeks Verlag.
- Ingwersen, P. (1992). *Information retrieval interaction*. London, England: Taylor Graham.
- Jacob, E. K. (1994). Classification and crossdisciplinary communication: Breaching the boundaries imposed by classificatory structure. In H. Albrechtsen & S. Oernager (Eds.), *Knowledge organization and quality management* (Proceedings of the third international ISKO conference) (pp. 101-108). Frankfurt/Main, Germany: Indeks Verlag.
- Jacob, E. K., & Albrechtsen, H. (1997). Constructing reality, : The role of dialogue in the development of classificatory structures. In I. C. McIlwaine (Ed.), *Knowledge organization for information retrieval* (Proceedings of the sixth international study conference on classification research) (pp. 42-50). The Hague, Netherlands: International Federation for Information and Documentation.
- Kirk, S. A., & Kutchins, H. (1992). *The selling of DSM: The rhetoric of science in psychiatry*.

**[Reference]**

- New York: A. de Gruyter.
- Nardi, B., & O'Day, V (1996). Intelligent agents: What we learned in the library Libri, 46(2), 59-88.
- Pejtersen, A. M. (1980). Design of a classification scheme for fiction based on an analysis of actual user-librarian communication and use of the scheme for control of librarians' search strategies. In O. Harbo & L. Kajberg (Eds.), *Theory and application of information research* (Proceedings of the second international research forum on information science) (pp. 146-159). London, England: Mansell.
- Pejtersen, A. M. (1994). A new approach to design of document retrieval and indexing systems for OPAC users. In D. I. Raitt & B. Jeapes (Eds.), *Online information 93* (17th international online meeting proceedings) (pp. 273-290). Medford, NJ: Information Today.
- Schmidt, K., & Simone, C. (1996). *Coordination mechanisms: Towards a conceptual foun*

**[Reference]**

dation of CSCW systems design. Computer supported cooperative work. The Journal of Collaborative Computing, 5(2-3), 155-200.

Star, S. L. (1989). The structure of ill-structured solutions: Boundary objects and heterogeneous distributed problem solving. In L. Gasser & M. N. Huhns (Eds.), Distributed artificial intelligence (pp. 37-54). London, England: Pitman.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, "translations" and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 190739. Social Studies of Science, 19, 387-420.

Star, S. L., & Strauss, A. (1999). Layers of silence, arenas of voice: The ecology of visible and invisible work. Computer supported cooperative work. Journal of Collaborative Computing, 8, 9-30.

Tang, R., & Solomon, P. (1998). Toward an understanding of the dynamics of relevance judgment: An analysis of one person's search behavior. Information Processing Management, 34, 237-256.

#### **[Reference]**

Thorhauge, J., & Segbert, M. (1997). Public libraries and the information society. Luxembourg:

Office for Official Publications of the European Communities. Wanting, B. (1984). How do children

ask questions in children's libraries?: Concepts of visual and auditory perception and language expression. Social Science Information Studies, 4, 217-234.

#### **[Reference]**

##### **ADDITIONAL REFERENCE**

Pejtersen, A. M.; Albrechtsen, H.; Lundgren, L.; Sandelin, R.; & Valtonen, R. (1996). Subject access to Scandinavian fiction literature: Index methods and OPAC development. Copenhagen: Nordic Council of Ministers (Temanord).

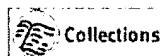
---

Reproduced with permission of the copyright owner. Further reproduction or distribution is prohibited without permission.

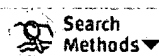
ProQuest

[Return to NPL Web Page](#)[Text Version](#)

English

[?Help](#)

Collections



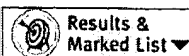
Search Methods



Topic Finder



Browse Lists



Results &amp; Marked List



Search Guide

Searching collections: All Collections

Article Display

[Email Article](#)

Article 80 of 82

[Publisher Info.](#)[Print Article](#)☐ Mark article

Article format: Full Text

[Save Link](#)

Saves this document as a Durable Link under "Results-Marked List"

## **DIGITIZATION EFFORTS AIM TO EASE PROCESS OF DATA RETRIEVAL**

*Chicago Tribune*; Chicago, Ill.; Aug 23, 1998; Andrew Zajac. Andrew Zajac covers technology companies for the Tribune.

**Sub Title:** [CHICAGOLAND FINAL Edition]

**Column Name:** BINARY BEAT.

**Start Page:** 5

**ISSN:** 10856706

**Subject Terms:** Digital imaging

**Companies:** UMI

**Abstract:**

*If Skokie-based Bell & Howell's UMI subsidiary and other information retrieval services have their way, these awkward encounters with ribbons of polyester film will be a thing of the past, replaced by digital files of text and graphics served up via the Internet.*

*Since June, UMI--once known as University Microfilms International--has been digitizing its vast archival holdings of newspapers, periodicals, research collections and scholarly dissertations with the aim of creating the world's largest on-line archive within five years. "This is a way of getting better access and better selectivity," said Dan Arbour, UMI's vice president for marketing.*

*Dubbed the Digital Vault Initiative, UMI's plan to place 5.5 billion pages of information on the **Internet** is part of the quickening pace of digitization of **library** and research **archives**.*

**Full Text:**

*Copyright Chicago Tribune Co. Aug 23, 1998*

**GUEST COLUMN.**

Click, whir, fumble. Clank, clank. Jiggle. Whoosh.

That is the onomatopoetic snapshot of frustrated scholarship, the sounds of loading, focusing, paying for, refocusing and printing out information from a microfilm reader.

If Skokie-based Bell & Howell's UMI subsidiary and other information retrieval services have their way, these awkward encounters with ribbons of polyester film will be a thing of the past, replaced by digital files of text and graphics served up via the Internet.

Since June, UMI--once known as University Microfilms International--has been digitizing its vast archival holdings of newspapers, periodicals, research collections and scholarly dissertations with the aim of creating the world's largest on-line archive within five years. "This is a way of getting better access and better selectivity," said Dan Arbour, UMI's vice president for marketing.

Dubbed the Digital Vault Initiative, UMI's plan to place 5.5 billion pages of information on the **Internet** is part of the quickening pace of digitization of **library** and research **archives**.

UMI has budgeted \$2 million this year for the digitization effort and is likely to spend similar sums for the next three to five years, Arbour said.



The project employs a dozen scanners that photograph microfilm data and convert it to digital images.

Each scanner runs 24 hours a day and is capable of creating 90 to 100 digital images per minute, according to Jeff Moyer, director of academic publishing.

UMI already has a limited amount of material from the last 10 years in digital format. The company's goal is to have complete editions of its 50 most popular periodicals, along with its 20 million-page collection of early English books produced from 1465 through 1700, available on-line by the second quarter of 1999.

Digitizing offers a couple of obvious advantages in addition to liberation from mechanical readers: easy, widespread distribution and flexibility.

"You can do in minutes what used to take days and in days what used to take months," said Remmel Nunn, publisher of Primary Source Media, a division of International Thomson, which plans to put up 7 million digitized pages of rare material pertaining to humanities research by 2000.

Information suppliers see revenue opportunities in the increased flexibility of a computerized archive. "We're going to build value-added slices of information," Arbour said. "We'll be adding bibliographies and offering the ability to select packages of information."

Primary Source, headquartered in Woodbridge, Conn., plans an offering that matches the Times of London Literary Supplement's book reviews.

Another significant digitization of scholarly material is the 7-year-old Literature On-line effort by microfilm publisher Chadwick-Healy of Alexandria, Va. This \$10 million-plus project will encompass 10 million pages and more than a millennium of English-language lit.

Perhaps the highest profile digital conversion under way is the Library of Congress' \$60 million National Digital Library Program, which aims to put up to 5 million sound recordings, rare documents, video clips and other items from its vast holdings on-line by 2001.

The Library of Congress is focusing its digitization program on unpublished, one-of-a-kind material, like original musical recordings and government-sponsored photographic collections, according to Bob Zich, director of electronic programs for the digital library project.

In an associated effort, the library is parceling out chunks of a \$2 million grant from Ameritech Corp. to digitize unique pieces of local and regional library archives.

Despite the advantages of on-line conversions, libraries and archives are unlikely to completely ditch microfilm.

For one thing, it remains an unrivaled medium for maintaining master copies of material. It stores compactly and has a shelf life of 500 years.

While an improvement, digitization is hardly an administrative or fiscal panacea.

"Say you're on a (computer) network and the network goes down. You don't have access. You're going to have to maintain some alternative forms of media," said Roberta Webb, director of the Chicago Public Library's general information services division.

In addition, merely digitizing doesn't automatically make a text searchable, leaving the user dependent on indexes from the original work. Searchability requires an extra layer of conversion, and expense.

Only the most popular portion of UMI's holdings will have this capability. But the Library of Congress' on-line archive is searchable.

---

Reproduced with permission of the copyright owner. Further reproduction or distribution is prohibited without permission.



Collections



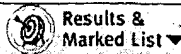
Search Methods



Topic Finder



Browse Lists



Results &amp; Marked List



Search Guide

Searching collections: All Collections

Article Display

[Email Article](#)

Article 144 of 151

[Publisher Info.](#)[Print Article](#)☐ Mark article

Article format: Text+Graphics

[Save Link](#)

Saves this document as a Durable Link under "Results-Marked List"

## The Internet Public Library: An intellectual history

*Library Hi Tech*; Bradford; 1998; [Joseph Janes](#);

**Volume:** 16

**Issue:** 2

**Start Page:** 55

**ISSN:** 07378831

### Full Text:

*Copyright MCB UP Limited (MCB) 1998*

#### [Headnote]

The ideas and work behind the Internet Public Library are described, from the original conception and initial set of projects through to current endeavors. Emphasis is given to questions asked about the nature of librarianship in a digital environment and projects and services that attempt to answer those questions.

### THE ORIGIN

This article will describe some of the ideas, decisions, and discussions behind the Internet Public Library (IPL), the impact they had, and how they turned out on the Web ( < <http://www.ipl.org/> > ).

The "developmental" period of the IPL, roughly from December 1994 to April 1995, can serve as a useful framework for discussing these ideas and decisions, but I will also lay out projects and ideas that followed.

### The Idea

In the fall of 1994, it was becoming increasingly clear that the Internet was going to have as significant an impact on libraries and librarians as on the wider world. Lou Rosenfeld and I had taught two successful courses in which we had had our students build detailed and extensive guides to Internet-based resources in specific subject areas. We had started to build gophers and learn about how they would work.

And then the Web happened. I can remember the first time I ever saw Mosaic; I can't say that I was thunderstruck by its potential to change the world, but it did look interesting and I wanted to learn more and play around with it.

Sometime in September 1994, the phrase "Internet Public Library" entered my consciousness. For several years, I had taught a seminar on the impact of information technology; each year, the theme had been different. It had been quite some time since I had last taught it. But now I was scheduled to do it again in the Winter 1995 term. I was planning to repeat the last theme, which was on intellectual property and technology. However, I wasn't too happy about the idea, because the last one hadn't been that successful.

It occurred to me that I might use this "Internet Public Library" idea as the theme for this course. I talked with several friends and colleagues and asked their opinions. They all said it sounded like a great idea but a lot of work. They were more correct than they knew, but I was encouraged and so went forward. I developed a prospectus for the course, which I circulated to the students in the then-School of Information and Library Studies (now School of Information) at the University of Michigan. Some relevant excerpts from that prospectus are reproduced in sidebar 1.

From the very beginning, the project was motivated by one central question: What does librarianship have to say to the network environment and vice versa? That question proved (then and now) both provocative and attractive. I asked students to submit statements of interest to join the class; I received over 50 statements, and selected a group of 35 students.

That group was quite diverse. Their backgrounds ranged from computer scientist and librarian to writer, editor, lexicographer, substitute teacher, IBM marketing rep, medical researcher, art historian, desktop video producer, human factors researcher, and medievalist. That breadth of experience came in handy, and the power of the central idea which brought them together focused their collective work.

In December 1994, in an e-mail message to the group, I laid out some further thoughts, based on a brief meeting and discussion we had just had:

#### My Current Ideas

are sketchy, but I think we need to create an entity that people can recognize both as a library and as a "true" Internet institution. That's a tightrope, but one worth walking. More specifically, I think we might publish a journal [with related content], provide a service helping librarians identify really neat Internet sources for use in their daily lives, and other things. The only thing I will insist on is a story hour; but how that might work I have no idea. Other people will probably propose other specific projects.

#### The Big Issues

that we need to resolve quickly are:

- \* the mission statement (not carved in stone, but helpful in deciding what we will and will not do),
- \* structure (of the overall project and of the groups that will do it; both administrative and functional), and
- \* evaluation methodology (how will you all be graded? A final exam, perhaps?).

Several important ideas are imbedded here. Most of the ideas in that first paragraph don't look much like what goes on in most libraries, but they seemed at the time (and still do) to be important aspects of exploring librarianship in this environment. We never did publish a journal or do collection development/profiling for other libraries, but we have returned to these ideas many times over the years, in thinking not only about services we might provide but also ways in which we might generate revenue. The story hour dictum was, in retrospect, really about making sure that whatever we did, it involved working with children in a meaningful way. In those days, there was almost no presence of children on the Net or resources for them, and this was long before the consuming concern about inappropriate or indecent material.

Those big issues have remained big. The initial mission statement has been revised only once, and has helped enormously in guiding work and thinking. Structure also helped, not only administratively but also in doing the work. Evaluation has been difficult and important. Pedagogically, evaluating students doing groundbreaking, original work within the same framework as grading people doing online searches and term papers has been an enormous challenge.

That same message started with the five goals I had outlined for the class:

- \* Finish the course more excited than when you started.
- \* Do work of which you and the entire group will be proud.
- \* Everybody (including me) learn a lot.
- \* Everybody (including me) have fun doing it.
- \* Everybody (including me?) get great jobs at the end.

This was an attempt to set the tone for the work, to help create an environment where people could explore, apply new technologies, stretch themselves and still have fun, yet work on something that would be meaningful and real. In general, I think we succeeded on all these counts.

#### THAT FIRST DAY

Our first meeting was on Saturday 7 January 1996. We met then to allow ourselves several hours in which to

define precisely what we wanted to do, who wanted to do what, our mission statement, and how to proceed.

## Mission and Goals

The mission statement came first. After some discussion, we adopted the following:

The mission of our Internet Public Library is to:

- \* provide services and information which enhance the value of the Internet to its ever-expanding and varied community of users;
- \* work to broaden, diversify, and educate that community; and
- \* communicate its creators' vision of the unique roles of library culture and traditions on the Internet.

This statement conveys our collective notion of what it meant to be a "library" in this chaotic, dynamic, placeless domain.<sup>1</sup> The first clause is fairly standard; the second betrays our desire to help people to understand the value of the Net and get more people on it. The final clause generated much discussion, especially around wording. We wanted people on the Net to know about the value of librarianship, what librarians know, and what they can do in this new environment.

The mission statement has since been changed. The newest statement, as adopted in July 1996, can be seen in sidebar 2. This is clearly a more elaborate and encompassing statement, laying out more precisely and fully our point of view and the perceived reach and scope of our work. It perpetuates the point of view and emphases of the original statement, though, and has served its purpose well.

## INITIAL GROUPS

After reaching consensus on the mission statement, small groups brainstormed on what exactly the IPL should be. These groups reported back and the entire class voted on what they thought were the highest priorities. The top six were:

- \* reference,
- \* architecture/interfaces/design,
- \* services for librarians/information professionals/ schools,
- \* bibliographic instruction/usereducation/information literacy/outreach/access,
- \* youth services, and
- \* public relations/development/legacy.

There are some interesting omissions here. Functions that are labeled "technical services" in libraries-collection development, organization of resources-were mentioned but received little support. These functions happened (within the rubric of reference), but this group did not see them as high priorities at the beginning. Possibilities that were mentioned but did not make the list included government resources, real-time interaction with humans, publication, art, community information/bulletin, an "Internet Advisory Conduit," a "900 Library" (fee-for-service research, another recurring idea), and others.

These groups then met to outline what they wanted to try to accomplish and how, roles within the group, and logistical matters. Several important decisions were made during those meetings and in the weeks that followed.

## Technology

Perhaps the most important decisions were made around the technology group (led by Nigel Kerr), which dealt with architecture, design, and related issues. Their summary of that first meeting captures those decisions well:

Goals include:

- \* unifying our IPL-technically, intellectually, for users; work closely with other groups;
- \* acquiring and managing server(s)/infrastructure; and
- \* providing opportunities for class members to learn about architecture and technical issues (Unix system

administration, HTML, interface design, listserv moderating).

In collaborating with other groups, we want to listen to their needs, use that info as input to our design and decision process, and inform them of what we think is/is not technically possible and feasible.

In other words, this group was not going to dictate what was possible or desirable, but rather chose to distribute themselves and act as liaisons, consultants, and advisors to the other groups. Those functional groups, such as reference and youth, were told to think first about what they wanted to do, without regard for technological capability or implementation, and the architecture group would see what could be done. This proved vital to the success of the project-it allowed for the free flow of ideas and creativity, which then could be tempered and adjusted as necessary based on the available technology.

Getting a server was an obvious initial hurdle. When the project began, we neither had one nor knew where we might get one. We were fortunate in this regard; Lee Liming, the school's technology administrator at the time, met with the architecture group, heard our ideas, and volunteered to donate a spare Sparc 20 server for our use. We are still using this server (with others) to this day; without it there would not have been an IPL at all.

The educational aspect was also important; people in all the groups learned what they needed to know in order to accomplish their projects. Rather than adopting technological matters as their sole province, the architecture group saw their role as facilitating other people, learning what they needed, and contributing as necessary.

## Reference

The reference group, led by David Carter, Nettie Lagace, Sara Ryan, and Schelle Simcox, emerged with the following goals (excerpted from an early report):

1. Create a virtual reference collection (a hotlist, but also a searchable database)
  - a. In this regard, collection development will be a big issue. We plan to divide ourselves into subject specialties, although not everyone may be able to specialize in a subject s/he likes (like the real world?)
  - b. We are also seriously considering classifying our collection along lines similar to those already used by M-Link.
  - c. What kind of format will the reference collection be? Comprehensive or selective?
  - d. How will we judge the authority of our sources?-identification, dates/updates, contact person, where source is located....
2. Provide online reference service, both via e-mail and in real-time.
3. Design the database structure. We want a product that can run on UNIX (so that it can be accessible from the Internet), but cost constraints may hamper this wish. The database must be able to handle big, ever-growing data files, and accommodate Boolean searching as well as field restrictions.
4. Develop a manual for online reference.
5. Collect usage statistics regarding the user community, types of questions asked, etc.

This prescient list of goals is remarkable in several regards: it predates almost all thinking and discussion in the professional literature about what "reference" means in a networked environment, it identifies a number of crucial issues that have continued to be discussed not only at the IPL but on the Net in general, and it emerges in almost equal measure from thinking about applying traditional librarianship to the Net (how to do collection development here) and the other way around (how to use the Net to provide reference service).

The original intent of the ready reference collection was to serve as a resource for answering reference questions. As it turned out, our collections (ready reference and beyond) have been our most popular and valued resources, and the actual reference-question-- answering service has taken place on a much smaller scale. This work is, as all librarians know, timeconsuming and labor-intensive, while collections, once built and maintained, are available for use by countless people. However, this kind of personalized service is still needed in a world of "virtual," digital collections-in fact, it is even more necessary in the distintermediated and incoherent Internet.

Important collection-oriented questions were and are asked here about format, selectivity, quality (but not, interestingly, about balance-not all points of view on issues have always been available in networked resources), access, and so on. Organization is also raised; M-Link, a service of the University of Michigan Library, had a gopher at this time, which was among the first organized along library principles. Sue Davidsen, who designed that gopher and who now heads the Michigan Electronic Library,<sup>2</sup> saw that opportunity and showed the way for many

projects (including, for example, the Argus Clearinghouse<sup>3</sup>) in applying librarianship to Internet-based resources.

Although the original ready reference collection was encoded in static HTML pages, it has since been moved to databases, per this original vision. At present, it and all other IPL collections are stored as FileMaker databases. Some are served out periodically as static pages; others (including the Online Texts collection) are dynamically drawn directly from the databases on the fly.

### Online Reference Help

The seemingly innocent fourth item, "Develop a manual for online reference," is in fact the precursor for a great deal of experimentation and exploration of what it means to do reference in this world. As our work has progressed, this has gotten somewhat easier, due to the increasing number and quality of network-based resources, the improved functionality of search engines and directories, and our growing experience. But the question of how best to integrate the use of the Internet and its resources into reference librarianship is an ongoing and continually challenging one.

At the beginning, there was no way to know, for example, how many questions we would receive. There was a continuing, half-joking notion of the 500 questions per day that we might get and how we would deal with them. Discussions about how many staff we would need and how we would solicit participation from librarians and subject experts followed.

Technological issues were important as well. How would we take questions, and what would happen to them? There was no enthusiasm for receiving e-mail messages in personal mailboxes, and it was obvious that guidelines would be needed to help people formulate questions so that they could be answered. Despite the desire to do real-time reference, it seemed obvious that using e-mail as the medium made sense, but that meant that reference interviews would be difficult or at least asynchronous and therefore time-consuming. An initial solution was the use of HyperNews, a product of the National Center for Supercomputing Applications, which allowed Web-based conferencing. This was inadequate, and within a few months, Michael McClellan, a Ph. D. candidate in computer science and member of the first class, wrote a new software package called QRC for handling reference questions; it is still in use at the IPL and several other libraries today.

To date, we still have never actively advertised the question-answering service, for fear of the volume of questions we might receive. Without advertising, and with the form buried a few levels down in the IPL site, we typically receive between 20 and 70 questions a day, which is often more than our crew of students and volunteer librarians can handle. As such, we have to reject some questions based on a loose quota system.

### Online Collections

When the IPL opened, ready reference was our only significant collection. Since that time, several more collections have been added: a collection designed for teenagers, a reading room for full texts, magazines and newspapers, and a directory of Web sites of nonprofit organizations. The youth division also added a small collection to supplement its offerings.

These collections have separate designs (although all share the overall IPL design dicta, discussed below), collection development procedures, selection policies, and organizational structures. There are areas of similarity, not surprisingly, but all these decisions have been revisited with each new collection.

For example, the original ready reference collection had a fairly simple structure: each resource was represented by its title, URL, a brief description, statement of authorship, and keywords (from an uncontrolled list), and placed in a category by subject. There was no attempt to create cataloging records, let alone full MARC records; there were too few resources and their nature did not seem to lend themselves to such depth of organization. However, when the online texts collection was organized, David Carter, the IPL head of collections, adopted Dewey classification—most of the resources at issue were from the 1920s and before (since copyright had lapsed and thus they were in the public domain), for which Dewey was intended to be used. As such, we have more comprehensive cataloging and records (though still not full MARC) for those records. We also use Dewey for the youth collection since most children use Dewey in schools and it is more likely to be familiar to them.

Original reference resources also have been created: POTUS (Presidents of the United States), on the American presidency; Stately Knowledge, on the U.S. states; Native American Authors; Say Hello to the World (introductions to 30 languages); and A+ Research and Writing (for high school and college students writing papers). These resources owe much to traditional reference publishing, but also take advantage of the hypertextual and multimedia environment, incorporating images and links to other relevant Web sites containing high-quality information.

In all cases, from the beginning, these collections and resources were based on personal or group interest and motivation. Bob Summers, the creator of POTUS, had a lifelong fascination with information about the presidents,

and used a summer-long independent study to compile information; organize it; structure, design, and build the site; find good external sites; and make the hundreds of decisions necessary to complete this excellent resource. Lorri Mon, who led the group that built Native American Authors and also compiled Say Hello to the World (while assisting in reference question administration), felt strongly about both topics, which reinforced the IPL's desire to reflect our users both within North America and around the world.

Again, this is librarianship in a new light. The collections work comes from a long and valuable tradition of collection development and maintenance. The original resources look more like publishing, but in all cases, the fact that they are designed to be in a library has strongly flavored their design, intent, and formulation.

## Youth

The youth division, led by Josie Parker, in its initial discussions, focused on these ideas:

- \* writing contest,
- \* FAQs from children to be answered by authors,
- \* publishing an interactive picture book,
- \* story hour,
- \* cool hot list,
- \* science project guide,
- \* kids in a MOO [Multi-user Object Oriented Environment], and
- \* book discussion group for older readers.

In general, this group was focusing on younger users (elementary school-aged) based on their interests and experiences. Many of these projects came to fruition: two writing contests drew several dozen entries from the United States and Europe; Ask the Author has children's book authors' answers to questions from kids and parents; several originally illustrated story books are available; and an original science project was put up. Some, which required more sustained and continual maintenance (the book discussion group) or more involved development work (interactive picture book) were not possible given the resources. The cool hot list evolved into the youth collection, now cataloged and organized using the Dewey Decimal Classification.

This list is another interesting mix of fairly traditional librarianship and Internet inventiveness and experimentation. The "required" story hour, familiar to most public libraries, is here, as are discussion groups, writing contests, and more; but the author FAQs, interactive book, and MOO participation are still amorphous-in the realm of speculation about serving young people in a networked, interconnected environment.

## Services to Librarians

The original mission statement committed us to share our lessons, not only with our patrons and the Internet world, but also with our professional colleagues. The Services to Librarians and Information Professionals group (led by Richard Truxall) worked on this. Their work broke into several large categories: help to librarians in getting connected to and using the Internet, examples of how libraries were using the Net, and professional resources. They wrote original documents on: the Internet, how to get connected, using it, and building resources (including telnet, Web, gopher, and Veronica); libraries using networks for their work; and Net resources from professional organizations. They also wanted to create a calendar of meetings, conferences, and events and foster a mentorship program to help librarians new to the network connect with more experienced colleagues.

In practice, these were successful, but maintenance again became a problem, and most of the ongoing nonWeb resources fell away over the years. This formed the nucleus for another set of ideas, though. Imagine, for example, the following consideration, taken from a 1996 proposal. By simply making the IPL accessible via the World Wide Web (WWW), we can reach effectively those people who use the Internet on their own. For the most part, these are people who own their own computers and modems. However, we noted that there is a large segment of the population who cannot afford direct Internet access, whose only connection to the Internet is through schools and public libraries.

In order to reach this population, the IPL proposed to work closely with other libraries (public libraries, school libraries, academic libraries) to help them effectively assist their patrons in using Internet resources. Doing this would involve 1) training their staff members to be Internet librarians, 2) providing our expertise to help solve their local problems, and 3) promoting the IPL as the initial place for their patrons to visit. As part of this, we proposed

developing a standard methodology which each library could present local information resources in conjunction with the IPL as a seamless whole.

Although many libraries have Internet connections, the comments we hear through professional channels and personal contact indicate that many of them have not yet integrated the Internet into their work. Many public libraries are organized into cooperatives or served by regional organizations expressly designed to support their members in dealing with new technology. However, the quality of this support varies widely and many libraries are on their own. We proposed to work with such agencies to help them help their own members, and with isolated libraries that have few other resources.

Certain realities have evolved from the 1996 proposal-for IPL specifically and librarians generally. We are not in the business of teaching people how to connect to the Internet and use its basic tools. Many other organizations can do better jobs of that than we can. What we can do best is to help librarians build upon their own professional training and apply their skills and experience in novel ways. We have found that the techniques and tools of professional librarianship translate well from the world of print to the world of digital information. We can build upon our experience to ease the transition for others, and to help individual libraries solve their own unique problems in this regard.

The IPL strongly resembles many other initiatives in which libraries and librarians have combined their time and resources to improve or expand the services they can offer to their users. Interlibrary loan, reference referral, and shared cataloging are simple examples; the use of network technologies can enable even more encompassing ideas.

#### Education and Outreach

Another group, led by Louise Alcorn, was interested in issues of education and user outreach-important functions in any library, and especially in one that would almost never meet its users in person. This group focused on making the Internet easier to use and understand, and increasing general knowledge of it and participation in it. This group worked:

- \* to provide pathways to access points and equipment for potential Internet users and to educate current and future users about techniques needed to use the Internet;
- \* to ensure access to the IPL by a variety of users, employing different platforms;
- \* to interest non-users in the "beauty" (and value) of the Internet;
- \* to learn the computing and communication needs of various audiences; and
- \* to provide navigation tools to the IPL, and, through the IPL, to the WWW.

Much of this reflected the thinking of the project as a whole; in particular the "different platform" issue, as discussed below in the section on look and feel, was a general concern. The rest, though, comes directly from traditional notions of librarianship: equality and ease of access to information and resources, knowledge of those resources, and ability to use them. One aspect of this, navigation, was implemented by building a directory-an alphabetical listing of resources and areas of the IPL. The directory was available when the IPL opened, but became a maintenance problem as the library grew and expanded. The directory, though, was a precursor of similar devices common in large Web sites today, particularly the site map.

In practice, these were (and are) enormous challenges, well beyond our small capacities at the time. The IPL has always remained committed to these ideas, but has not been able to achieve as much as the vision of this group would have required. Much of this invaluable work goes on now in libraries of all kinds, all over the country, and appropriately so. Today, libraries and librarians in direct contact with their users and communities can far more effectively put them in touch with the Net and help them to use Web resources than the IPL could ever hope to do.

One other aspect of the IPL grew out of the outreach group. One member of that group, Kendra Frost, wanted to extend the benefits of the project to the Museum of African American History in Detroit. At the time the museum had no presence or even awareness of the Internet, but she felt this was a chance for them to tell their story, share their resources, and learn about the potential of the Net. With their permission, she built a few pages describing the museum and its work, and incorporating images of a few items from its collections. It looked much like the sort of small exhibit one often finds in libraries, and so we built an exhibit hall to house it. Since that time, several exhibits have entered the exhibit hall, each allowing its creators to explore design and work with images, multimedia, sound, or other more exotic technologies. Exhibits have ranged from photographs of trains and lighthouses to the story of Detroit jazz, ancient Egyptian forgeries, Pueblo pottery, dinosaurs, and the life story of a woman and her family (Grammy Mirk).



## Public Relations

Finally, there was the public relations group, led by Maria Bonn and Bradley Taylor. In early 1995, the Web was still a young and comparatively unpopulated place. Students who were interested in or had experience with public relations undertook this responsibility. Their original mandate, which included documentation of our work, fundraising, and development efforts, was really beyond the scope of what was possible; however, the public relations effort was more successful than any of us could have imagined.

Their priorities, from the first meeting, were to

### 1) Implement publicity and promotion by

- \* Saturating the Internet with ongoing press releases,
- \* generate local and national media exposure, and
- \* Publicizing IPL via ALA, SAA, SLA, etc., newsletters and conferences;

### 2) Facilitate internal communication by

- \* Publishing an in-house newsletter and
- \* Providing liaison between groups;

### 3) Develop documentation by

- \* Saving everything,
- \* Producing a video record,
- \* Compiling history of process from documentation at end of semester, and
- \* Writing a final article; and

### 4) Implement fundraising/development by

- \* Seeking university support and
- \* Looking for corporate sponsorships.

The first of these was accomplished. (These were still the days when an e-mail message posted to the right newsgroups and listservs could have a large impact.) The press release, sent to the world (but targeted to the library world), was extremely well written, generated great interest and enthusiasm, and raised the stakes dramatically for the work as a whole. Excerpts from the text of that release are shown in sidebar 3.

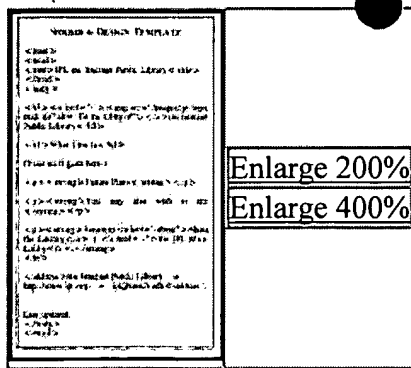
The response was overwhelming. So many requests for information were received (going to individuals' e-mail boxes) that we had to set up a listserv to keep people informed on the work. Subsequent messages encouraged people to subscribe to that group. The press release was sent in February; by the time the IPL opened in mid-March, over 3,000 people from around the world had subscribed.

Now, several thousand people were waiting to see what we would come up with. This marked the transition from a class project that would be up for a few weeks, to a real product, facing a global audience that would want to see it, use it, and perhaps even depend on it. As the number of people on the listserv grew, so did the pressure to produce a high-quality product by the deadline.

Never underestimate the power of a well-written press release.

## LOOK AND FEEL: DESIGN AND DICTA

One of the overarching themes of the early discussion about what the IPL should and would be was making it a "place" where people could explore, relax, read, and so on. The metaphor of a physical library was so strong it was often unspoken; it was taken for granted, and people would talk about the planned IPL in much the same way they would a physical library. Metaphor turned out to be one of our best friends—the more we were able to discuss the IPL in terms of rooms and services and places found in libraries, the easier the work got and the more ideas we generated.



SIDEBAR 4: DESIGN TEMPLATE

The importance of a sense of place in an inherently placeless environment should not be underestimated. It cuts to the heart of what a "library" is and what it means in a distributed world: a refuge, a stable island in a sea of chaos, an organizing force. It also conveys continuity and durability.

This formed the backdrop for conversations about what the IPL would actually look like. These conversations reflected the desire (in fact, more like a need on our own part) for stability and place, but also a recognition that many people initially would be coming via slow connectivity, using non-graphical browsers such as Lynx, or would have image loading turned off to increase navigation speed. We were particularly thinking here of people with low-bandwidth connections, especially internationally. It came as a pleasant surprise, then, when we received many e-mail messages from sight-impaired users praising our text-rich design. Software that vocalizes text on the screen does not work well with images, and so our site was ideal for their use. The recent federal ruling concerning ADA compliance of Web design confirms our vision in this regard.

There were numerous impassioned discussions about design and consistency during the building period. There was strong sentiment both for basic consistency across the library and for separate divisions to be able to design for specific categories of users and to implement other ideas. The middle ground here was to let all groups experiment for several weeks, trying out different designs, and then combine the best elements to create a single look and feel for the library. This approach worked (though not without continuing debate), and produced the first set of design dicta and template. That template is shown in sidebar 4.

The design template is fairly simple, but therein lies its power. The point of the template and the dicta behind it was to provide "wallpaper and carpeting," a basic look for all IPL pages. This freed people from thinking about design to think more about content. The page was designed to be simple, quick and easy to load, and unobtrusive, yet provide structure and content.

The logo and name of the library appear at the top of each page, so it is immediately clear that you're in the IPL, and equally clear when you've left. All pages start with a first-level heading and end with a standard footer incorporating the name of the library, its URL, its e-mail address, and the date the page was last updated. Just above that are links back up through the hierarchy of the library sending people, level by level, up to the division home page and the library's home page (called the main lobby, to reinforce the building metaphor). Some pages also had a "You may also wish to see" tag, suggesting related IPL sites.

All IPL pages had to comply with these dicta, with only a few exceptions, including the home page and divisional main pages (reference, services, education), which could incorporate graphics. Each of these, though, had a text-only version in addition to the graphical version. The youth division (and later, exhibits) was granted a blanket exemption; they made a compelling argument that children required larger text, color, more graphics, and big buttons to click on; their designs were based on the dicta when applicable, but also reflected these additional needs.

They also had a larger, color version of the IPL logo, which also was designed to be simple, easy to load, and distinctive. The small logo, which appears on each page of the IPL, is a grayscale .gif file of only 1K. Over time, other versions of the logo have been designed for various projects and resources; it has proved to be quite flexible. These other versions include color but maintain the same structure and design to permit individuality within consistency.

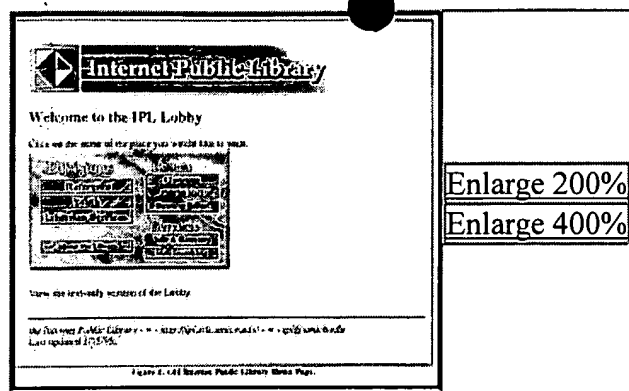


Figure 1:

The main lobby went through several designs. The first listed all IPL divisions and resources with lengthy descriptions. These descriptions (apparently more for our benefit than anything else) vanished but the architectural device above it remained.

In retrospect, this was obviously intended to convey stability. Shortly before we opened, it was replaced by the design shown in figure 1, similarly architectural but meant to resemble a plaque.

This design persisted for over two years. Many discussions arose after that time about changing it, but it didn't happen until mid- 1997, when an updated home page, designed by Robert Mann and shown in figure 2, was implemented.

Discussions about design, especially the front page, have often verged on the emotional. There seems to be great depth of feeling about look and feel, which is not surprising. To the world, this conveys who we are-our public face and the first impression people take away. This poses a multiple challenge: we need to be interesting and inviting (which might mean changing designs and using more graphics) but we also need to emphasize consistency (which would argue for few changes). Inertia also plays a role here; but we have incorporated several graphical and design changes while concentrating on adding new resources and enhancing existing ones.

## BEYOND THE BEGINNING: WHAT WE HAVE LEARNED

As of this writing, the IPL is over three years old. We have been visited over 7.5 million times by users from over 130 countries, and answered almost 7,000 reference questions. We presently are averaging over 20,000 users per day, an average of one about every five seconds, 24 hours a day. Our collections now number over 17,000 items, including nearly 7,000 pointers to online texts, 2,600 reference resources, 2,300 serials, 2,000 newspapers, nearly 1,100 associations, and 1,000 items for young people.

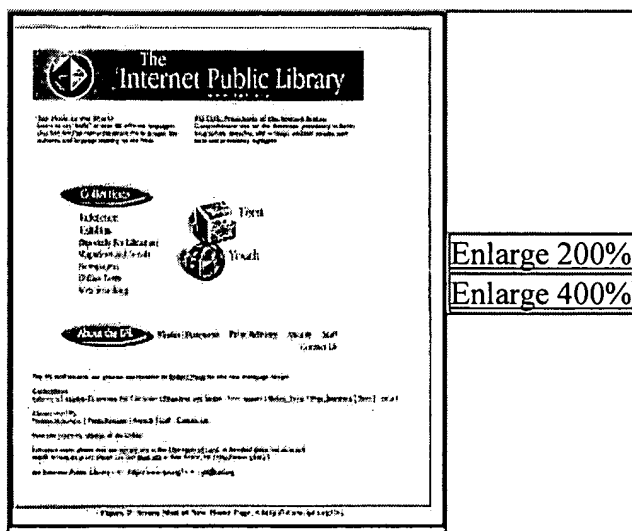


Figure 2:

We have learned several important lessons. Three of the more important ones are discussed below.

Librarianship works... almost all the time.

This can be restated: the IPL performs many functions that almost every other library does, and these are both

exactly the same and completely different. We answer reference questions, but we can't see people, read their facial expressions or body languages, or even interview them very well. We tell stories, but to children we never meet. We select, describe, and organize resources, but we don't catalog. Yet in all these instances, traditional librarianship has guided what we do and how we do it. In fact, when faced with a challenge or problem, we almost always explore approaches from the profession, and more often than not, they are helpful.

There also have been instances where that doesn't work. For example, we often have found that when we send e-mail to people who have asked us reference questions, to follow up or ask for more information, we get no response. Whether it's because they don't check e-mail often, don't care about their question that much, have lost interest, or no longer have e-mail, we don't know. There really isn't an analogy to that in the real world-sometimes patrons do drift away, but it's hard to imagine asking someone a question and that person just standing there, making no effort to answer at all. So what do we do? Answer as best we can? What if the patron truly doesn't care any more? Should we use our scarce resources to answer a question for a patron who has evaporated, as far as we can tell?

Most libraries attempt to build collections that will endure, and, of course, books and other physical carriers of information will persist, at least until they are stolen, weeded, or replaced by new editions or better resources. It is, in practice, difficult for us to know which "items" in our "collections" are even still there at any given point, whether they have changed, or what's happened to them. It certainly makes collection development (and maintenance) a challenge.

Technology is not the point... but this one is different.

Libraries and librarians have become masterful at incorporating new technologies and storage media into their work. Walk into almost any library, and you'll see not only books but magazines, newspapers, audio CDs, CD-ROMs, video cassettes, pamphlets, posters, art, and so on, as well as connections to digital resources from commercial vendors such as DIALOG and, of course, the Internet. What we always have known is that the medium is less important than the quality of information and that technology can help provide new and better kinds of access to information.

But, of course, those technologies are not the point. Most libraries have microforms of some flavors, but the incorporation of microfilm or fiche didn't fundamentally change librarianship. Neither did online resources or CD-ROMs. Frankly, neither will the simple presence of Internet-based resources. They are simply another set of potential aids in helping people find out more about their world and lives. They do raise some new and fascinating questions about the nature of publishing and authority and the value of editing, but just having access to the Net won't change librarianship or libraries.

On the other hand, effective use of the Net might change librarianship and libraries. There is a difference to this technology, as compared to those that have gone before. The Internet is more than just a new storage medium or search facility. The power of the Net, in librarianship as everywhere else, is its ability to make connections to people, organizations, ideas, and information. It could facilitate major change and a quantum leap forward in the quest to allow people to be more fully informed and aware.

The central problem of librarianship is to help get information out of one person's head and into another's. We have been doing this with books and indexes and catalogs and reference books because that's what we've had. When you add the ability to connect directly to people and organizations and communities who know about diverse subjects and can make information directly available, you open a new paradigm.

You can break it open even more widely when you connect librarians. Listservs such as STUMPERS-- L and Web4Lib are one thing, but consider the power of thousands of librarians, connected via the Internet, working together on collection development, readers' advisory, reference, storytelling, and all the rest, to serve millions of people on a daily basis. It staggers the imagination, and places librarianship smack in the middle of a revolution in information provision. The Internet won't be the end of libraries, as many have proclaimed-it could be the beginning of the enshrinement of librarianship and what it stands for as one of the most important, valuable, and respected professions.

The best way to learn it is to do it.

There is simply no way that any of us could have learned the lessons we have over the years if we hadn't been out there developing the Internet Public Library. As we continue to develop new capabilities, we continue to learn, both from applications that work well and from implementations that don't.

I said over and over in that first class that there was no way for that work to fail. I believed it then and still do. Regardless of how many people use the IPL, and what they think of it, the only way for the work to fail is if nobody learns from it.

The IPL always has been a vehicle for learning and trying new technologies. It can be thought of as a teaching and research library in the model of teaching and research hospitals-where people come to learn how to heal more effectively and try out new methods of treatment, all the time interacting with real patients and providing a real service. IPL is much the same; students come to learn about librarianship in the emerging information environment, librarians come to get new perspectives and ideas and continuing education,<sup>4</sup> and thousands of people find the Internet a hospitable source of information.

#### The IPL Today

And that is how the IPL is proceeding today. We have a small professional staff (1.75 FTE plus a large fraction of my time as director) to coordinate and provide training and continuity. The professional staff is supported by a class of students and other volunteer professional librarians. This class is part of the new Practical Engagement Program of the School of Information, which is designed to give students an opportunity to learn by doing through structured workshops engaging communities in the larger world. In our case, that community is not only our users but also professional colleagues who wish to work with us and participate in our projects.

That work (maintaining and expanding our collections; answering reference questions; providing stories, exhibits, and original resources) continues the tradition of librarianship: providing a sense of place, resources, and services to help people find information they want or need.

#### [Sidebar]

##### SIDEBAR 1: EXCERPTS FROM "INTERNET PUBLIC LIBRARY" PROSPECTUS

#### [Sidebar]

Much discussion and attention has recently been focused on the **building "digital libraries,"** including federations of **digital** and, presumably, other kinds of collections of documents with sophisticated search engines and techniques, user interfaces, and other features. The recent National Science Foundation-- sponsored grant competition has generated a great deal of interest in this area, and six large projects are underway (including one at Michigan). This effort will be somewhat different. We will explore the issues involved in the merger of networking and libraries by actually planning, building, and running a public library for the Internet community. Some of the larger questions raised by such a project include:

#### [Sidebar]

- \* To what extent do the kinds of functions traditionally seen in libraries apply to this setting?
- \* To what extent do the kinds of tools and resources traditionally seen on the Internet apply to this setting?
- \* What unique functions or features will be necessary or desirable in an Internet library?

#### [Sidebar]

Clearly, it would be undesirable to simply replicate a public library, function by function, job by job, process by process on the Internet. Similarly, the sorts of "libraries" one finds available over the Net today are little more than long lists of resources or categories of resources involving little intellectual effort or input from the library community. An Internet Public Library would be a true hybrid, taking the best from both worlds but also evolving its own particular features.

I would view each of those three words as equally important in conveying the intent of this project: Internet, Public, and Library. I think the combination of the three of them produces something quite different than any pair or individual might suggest.

Certainly, when planning any public library, one must take account of the public being served. Who is our "public"? Whoever they are, how will we best serve them? And how will we know if we are adequately serving them or not? Usage statistics? Formal feedback?

How is the library to be designed?

Will the library create and make available products of its own?

The primary objectives of this project are:

#### [Sidebar]

- \* to motivate the discussions about the role of libraries in the networked environment and networks in the library environment; and
- \* to provide students with a real experience in planning and executing a major- project with potential to serve a real user community in a meaningful way.

#### [Sidebar]

##### SIDEBAR 2: MISSION STATEMENT OF INTERNET PUBLIC LIBRARY

**[Sidebar]**

The Internet is a mess. Since nobody runs it, that's no surprise. There are a lot of interesting, worthwhile, and valuable resources out there-and a lot that are a complete waste of time. Over the last few hundred years, librarians have become skilled at finding the good stuff, organizing it, and making it easier for other people to find and use. Librarians also fight for important ideas like freedom of expression and thought, equality of access to information, and literacy.

The Internet Public Library is the first public library of the Internet. As librarians, we are committed to providing valuable services to that world. We do so for many reasons: to provide library services to the Internet community, to learn and teach what librarians have to contribute in a digital environment, to promote librarianship and the importance of libraries, and to share interesting ideas and techniques with other librarians.

Our mission directs us to:

**[Sidebar]**

- \* serve the public by finding, evaluating, selecting, organizing, describing, and creating quality information resources;

**[Sidebar]**

- \* develop and provide services for our community with an awareness of the different needs of young people;

**[Sidebar]**

- \* create a strong, coherent sense of place on the Internet, while ensuring that our library remains a useful and consistently innovative environment as well as fun and easy to use;

**[Sidebar]**

- \* work with others, especially other libraries and librarians, on projects which will help us all learn more about what does and does not work in this environment; and

**[Sidebar]**

- \* uphold the values important to librarians, in particular those expressed in the Library Bill of Rights.

**[Sidebar]**

We are committed to providing free services to the Internet community, in the greatest tradition of public libraries. However, we cannot sustain our library without a solid financial base. We are continually seeking enterprises that provide both service to our community and funding for our operations. We are always open to new ideas and partnerships.

**[Sidebar]**

**SIDEBAR 3: BOLD INITIATIVE HERALDS THE CREATION OF TOMORROW'S LIBRARY TODAY**

**[Sidebar]**

The University of Michigan School of Information and Library Studies proudly announces the advent of the Internet Public Library (IPL), an innovative, on-line, 24-hour public library designed to revolutionize the way the world thinks about library services. The Internet Public Library will offer an exciting version of the library of tomorrow as envisioned by many of the brightest talents in the field today.

With a stated mission to "provide services and information which enhance the value of the Internet to its ever expanding and varied community of users," IPL is prepared to provide essential library services to a target audience estimated to number 1/4 of the entire American population by the end of the century. Among the first services to appear will be an on-line reference division; a youth services division; a user education division; and professional services for librarians. A library without windows, walls, or even books, IPL will still provide a user friendly spot to turn to for questions on how to plan a family budget; learn more about the world of Internet-based resources; or even turn to for a story hour for children. Bringing the best features of the community library forward into a new technological environment, IPL seeks to challenge our thinking about new "communities" that will arise in the future and a broader range of services the library of tomorrow might provide.

**[Footnote]**

Acknowledgment: This article is dedicated with gratitude to all the people who have helped to make the Internet Public Library what it has been and continues to be. Your work has been a source of great pride, joy, and inspiration to me from the beginning.

**[Footnote]**

## NOTES

**[Footnote]**

1. A word on the first line: the word "our" appears instead of "the" consciously-we wanted to allow for the possibility that other "internet public libraries" would arise, so we could only speak for our own. In fact, there was a Web site calling itself the Internet Public Library before we started-it was a commercial site in Canada. I had found it, but neither communicated with them nor told the class about it, because I didn't want it to influence our work. I noticed several months later that the Canadian site had changed its name to "Cybrary," and as yet, no other IPLs have arisen.

**[Footnote]**

2. <<http://mel.lib.mi.us/>>. See Susanna L. Davidsen, "The Michigan Electronic Library," Library Hi Tech 15:3-4, consecutive issues 59-60 (1997): 101-106.

**[Footnote]**

3. < <http://www.clearinghouse.net/>>.

**[Footnote]**

4. We always are looking for volunteers to help us, in all areas of our work. Anyone who's interested in helping out for a few hours a week can send us an e-mail at [ipl@ipl.org](mailto:ipl@ipl.org).

**[Author note]**

Joseph Janes

**[Author note]**

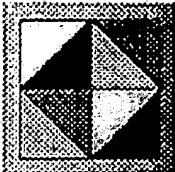
Janes is director, The Internet Public Library, School of Information, University of Michigan, Ann Arbor, Michigan. <[janes@umich.edu](mailto:janes@umich.edu)>.

**[Author note]**

Joseph Janes  
University of Michigan  
Ann Arbor, Michigan

---

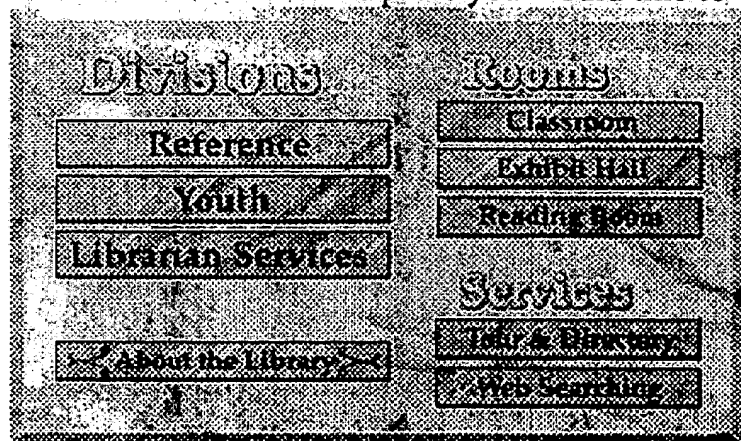
Reproduced with permission of the copyright owner. Further reproduction or distribution is prohibited without permission.



# Internet Public Library

## Welcome to the IPL Lobby

Click on the name of the place you would like to visit:



View the text-only version of the Lobby.

---

*the Internet Public Library - = - <http://ipl.sils.umich.edu/> - = - [ipl@umich.edu](mailto:ipl@umich.edu)*  
Last updated 3/15/95.

**Figure 1: Old Internet Public Library Home Page.**





# The Internet Public Library

www.ipl.org

## Say Hello to the World

Learn to say "hello" in over 30 different languages, plus find links to resources about the languages, the cultures, and language learning on the Web.

## POTUS: Presidents of the United States

Comprehensive site on the American presidency includes biographies, speeches and writings, election results, odd facts and presidency highlights.

### Collections

Reference

Exhibits

Especially For Librarians

Magazines and Serials

Newspapers

Online Texts

Web Searching



Teen

Youth

### About the IPL

Mission Statement | Press Releases | Awards | Staff  
Contact Us

The IPL staff extends our greatest appreciation to Robert Mann for this new frontpage design!

#### Collections

Reference | Exhibits | Especially For Librarians | Magazines and Serials | Newspapers | Online Texts | Web Searching | Teen | Youth

#### About the IPL

Mission Statement | Press Releases | Awards | Staff | Contact Us

View the text-only version of the Lobby.

European users: please visit our mirror site at the University of Lund, in Sweden! (<http://ipl.ub.lu.se/>)

North American users: please use our main site in Ann Arbor, MI. (<http://www.ipl.org/>)

the Internet Public Library - - <http://www.ipl.org/> - - [ipl@ipl.org](mailto:ipl@ipl.org)

Figure 2: Screen Shot of New Home Page <<http://www.ipl.org/>>.

## SIDEBAR 4: DESIGN TEMPLATE

```
<html>
<head>
<title>IPL the Internet Public Library</title>
</head>
<body>

<h3> <a href="/">  </a> the Internet
Public Library</h3>

<h1>What This Is</h1>

(Your stuff goes here.)

<p> <strong>Future Plans</strong> </p>

<p> <strong>You may also wish to see
</strong> </p>

<p> <strong>Return to <a href="/about"> About
the Library</a> | <a href="/"> the IPL Main
Lobby</a> . </strong>
<hr>

<address>the Internet Public Library - = -
http://www.ipl.org/ - = - ipl@umich.edu</address>

Last updated
</body>
</html>
```

ProQuest®

[Return to NPL Web Page](#)[Text Version](#)

English

[?Help](#)

Collections

Search Methods

Topic Finder

Browse Lists

Results &amp; Marked List

Search Guide

Searching collections: All Collections

Article Display

[Email Article](#)

Ⓢ Article 132 of 141 Ⓢ

[Publisher Info.](#)[Print Article](#)☐ Mark article

Article format: Full Text

[Save Link](#)

Saves this document as a Durable Link under "Results-Marked List"

## A guide through the Web's maze

*Businessline*; Islamabad; Jun 18, 1998;

NAICS:514199

Sub Title: [2]

Start Page: 1

Companies: [Yahoo](#) Ticker:YHOO NAICS:514199

### Abstract:

The company is able to make available all this information as it has been gathering, managing and analysing "multi-terabyte collections of information about **Internet sites**" ever since its founding in April 1996. (A terabyte is a million megabytes or 1,000,000,000,000 bytes). According to the company, its Web **archive** is already in excess of Eight Terabytes and it takes a new "snapshot" (using software robots) approximately every 30-60 days. [Alexa] is also donating its data - the company's "copy of the Web" - to an non- profit **Internet Archive** where it will be kept as a resource for scholars and historians to "help preserve our collective digital heritage." The name Alexa itself is derived from the famous ancient **library** at Alexandria. In fact, the company's **archive** is so comprehensive - "500,000 Sites and Growing" - that even when the real owners of a site have taken out information from their sites, the chances are that those pages still remain on Alexa's **archives** and accessible through the "**Archive of the Web**" button on the toolbar. When the user comes across the dreaded "404-Not Found" message (indicating a missing page) and Alexa has a copy of that page, the **Archive** button will change to blue. By clicking on the icon, the user can request a copy of that page. As a bonus, the Alexa toolbar also provides easy access to leading online reference tools such as the Encyclopaedia Britannica Online and Merriam-Webster's Dictionary and Thesaurus. Alexa is able to provide its service free to users by displaying small ads on the toolbar, when their users go online. Companies such as [Bank of America](#) and the Ziff-Davis publishing group are currently among the larger advertisers. Personally, I find Alexa's very useful. However, since the toolbar is only useful when logged on the **Internet**, it can be an irritant when you only want to use your browser to read files located on your PC's hard disk. Also, though Alexa's tool is generally "aware" of the important sites, it understandably does not cover all pages on the Internet.

### Full Text:

*Copyright Asia Intelligence Wire from FT Information Jun 18, 1998*

DESPITE the best attempts of search engines such as [Yahoo!](#), Excite and Infoseek, most Internet users would agree that the only "sure shot" way to locate information that would interest them on the Net is through personal recommendations of like-minded friends and associates. One company, the US-based Alexa Internet, is using this phenomenon to make Web-browsing a less chancy affair for users and even "change the way we use the Internet forever".

"When you walk on a path through the woods, you are benefiting from the exploration people have done before you- finding the best way up the mountain or down to the lake. Alexa tries to do the same thing for the Web," the company says.

"We all experience the gaps in navigating and finding information on the Internet as it is used today - frustrating keyword searches that turn up hundreds or thousands of Web pages and sites, very few of which are of any interest. What if we, as a community of users, could effortlessly pool our collective experience and add human

intelligence to navigation? What if we could fill in those gaps? It is a radical concept and it is our goal to make it real, and we invite you to join us in this effort," Alexa says at its Web site ([www.alexa.com](http://www.alexa.com)). To achieve this goal, Alexa has launched a different type of search engine which, it claims, is "the first Internet software product that learns from people." Unlike "normal" search engines such as @Yahoo!, users need not visit Alexa's Web site each time they want to search the Web. All they have to do is download a small piece of software - which is available for free - from the company's Web site. Once downloaded and installed, the software assumes the form of a small horizontal "toolbar," which sits on the bottom of the user's Web browser window. The tool then serves as an intelligent navigator that accompanies the user as he moves from site to site on the Web and "provides a continuous source of relevant recommendations of where to go next."

The Alexa toolbar is automatically triggered open whenever users switch on their Web browser. The tool looks at the site a user is currently viewing and suggests other pages by analysing where previous visitors to that site went next. As users visit a particular site, a list of related sites appear on the "Where to Go Next" portion on the Alexa toolbar. "For example: You need to find a new Internet Service Provider but you're not sure where to start. All you need to do is navigate to your existing provider's Web site, and then pop up the "Where to Go Next" menu to find a list of other Service Providers," the Alexa site explains.

How does Alexa come with the list of related sites? "Whenever your browser goes to a Web page, the Alexa Toolbar requests information about that page from Alexa's servers. We then record that an Alexa user has spent time at that site as a kind of vote - we do not know who passed through this site, but we know how many have passed this way. We can even learn from the fact that users visit several related sites as they browse the Web, and tend to spend more time at the sites that give them more of what they are looking for," the site's Frequently Asked Question (FAQ) section explains.

Alexa also learns from users by letting them explicitly make suggestions of "Where to Go Next" from a given Web page. Users can do this by selecting the "Add a link to this list" button on the Alexa Toolbar. "It's like letting people put up their own signs in the woods," the FAQ says.

The "Where You Are" button on the toolbar informs the user about who has registered the site he is currently visiting, the amount of traffic it receives, the speed of the site and the "freshness" of the information contained in it.

The company is able to make available all this information as it has been gathering, managing and analysing "multi-terabyte collections of information about Internet sites" ever since its founding in April 1996. (A terabyte is a million megabytes or 1,000,000,000,000 bytes). According to the company, its Web **archive** is already in excess of Eight Terabytes and it takes a new "snapshot" (using software robots) approximately every 30-60 days. Alexa is also donating its data - the company's "copy of the Web" - to a non-profit **Internet Archive** where it will be kept as a resource for scholars and historians to "help preserve our collective digital heritage." The name Alexa itself is derived from the famous ancient **library** at Alexandria. In fact, the company's **archive** is so comprehensive - "500,000 Sites and Growing" - that even when the real owners of a site have taken out information from their sites, the chances are that those pages still remain on Alexa's **archives** and accessible through the "**Archive of the Web**" button on the toolbar. When the user comes across the dreaded "404-Not Found" message (indicating a missing page) and Alexa has a copy of that page, the **Archive** button will change to blue. By clicking on the icon, the user can request a copy of that page. As a bonus, the Alexa toolbar also provides easy access to leading online reference tools such as the Encyclopaedia Britannica Online and Merriam-Webster's Dictionary and Thesaurus. Alexa is able to provide its service free to users by displaying small ads on the toolbar, when their users go online. Companies such as @Bank of America and the Ziff-Davis publishing group are currently among the larger advertisers. Personally, I find Alexa's very useful. However, since the toolbar is only useful when logged on the **Internet**, it can be an irritant when you only want to use your browser to read files located on your PC's hard disk. Also, though Alexa's tool is generally "aware" of the important sites, it understandably does not cover all pages on the Internet. After all, the number of sites is said to be doubling every six months and the number of Web pages has already crossed 300 million.

"Please download Alexa, use the service, vote on pages, recommend links, and help make our vision even better. It will change the Net for all of us," the Alexa site exhorts Net users. Despite the minor irritant pointed out above, I agree.

(The author, a freelance writer, can be reached at [arun.natarajan@cheerful.com](mailto:arun.natarajan@cheerful.com))

---

Reproduced with permission of the copyright owner. Further reproduction or distribution is prohibited without permission.